THE PHYSICS AND CHEMISTRY OF SOLIDS

AN INTERNATIONAL JOURNAL

VOLUME 13 1960

VOL. 13 1960

Editor-in-Chief HARVEY BROOKS *Cambridge, Mass.*

Associate Editors H. B. G. CASIMIR *Eindhoven* G. J. DIENES *Brookhaven* JACQUES FRIEDEL *Paris*

E. M. LIFSHITZ *Moscow*

Editorial Advisory Board

U.S.A.

E. N. ADAMS, Pittsburgh
D. ALPERT, Illinois
L. APKER, Schenectady
J. BARDEEN, Illinois
D. S. BILLINGTON, Oak Ridge
R. G. BRECKENRIDGE, Cleveland
LEO BREWER, Berkeley, Calif.
E. BURSTEIN, Penn.
J. A. BURTON, New Jersey
N. CABRERA, Virginia
G. C. DANIELSON, Iowa
I. ESTERMANN, Washington, D.C.
J. C. FISHER, Schenectady
F. HERMAN, Princeton
CONYERS HERRING, New Jersey
A. R. VON HIPPEL, Cambridge, Mass
J. H. HOLLOMON, Schenectady
W. V. HOUSTON, Houston
H. B. HUNTINGTON, Troy
H. M. JAMES, Purdue
JOHN G. KIRKWOOD, Yale
C. KITTEL, Berkeley, Calif.
W. KOHN, Pittsburgh
J. A. KRUMHANSL, Cleveland
A. W. LAWSON, Chicago
H. W. LEVERENZ, Princeton
O. T. MARZKE, Pittsburgh
B. T. MATTHIAS, New Jersey
L. R. MAXWELL, Silver Spring
JOSEPH E. MAYER, Chicago
E. W. MONTROLL, Maryland
DAVID PINES, Princeton
GEORGE T. RADO, Washington, D.C.
H. REISS, New Jersey
J. R. REITZ, Cleveland
A. ROSE, Princeton
O. C. SIMPSON, Argonne
R. SMOLUCHOWSKI, Pittsburgh
F. H. SPEDDING, Iowa
D. TURNBULL, Schenectady
J. H. VAN VLECK, Harvard
B. E. WARREN, Cambridge, Mass.
EUGENE P. WIGNER, Princeton
B. H. ZIMM, Schenectady

Canada

R. E. BURGESS, Vancouver
D. K. C. MACDONALD, Ottawa

United Kingdom

A. CHARLESBY, Shrivenham, Berks.
W. J. DUNNING, Bristol
F. C. FRANK, Bristol
H. K. HENISCH, Reading
H. JONES, London
J. W. MITCHELL, Bristol
A. B. PIPPARD, Cambridge
M. H. L. PRYCE, Bristol
R. A. SMITH, Malvern
K. W. H. STEVENS, Nottingham
F. C. TOMPKINS, London

France

E. F. BERTAUT, Grenoble
A. GUINIER, Orsay
LOUIS NÉEL, Grenoble
LOUIS WEIL, Grenoble

Germany

M. BORN, Bad Pyrmont
ULRICH DEHLINGER, Stuttgart
R. HILSCH, Göttingen
G. LEIBFRIED, Aachen
E. MOLLWO, Erlangen
H. PICK, Stuttgart
M. VON LAUE, Berlin
H. WELKER, Erlangen

Italy

F. G. FUMI, Pavia

Czechoslovakia

J. TAUC, Praha

Japan

TOSHINOSUKE MUTO, Tokyo

The Netherlands

A. MICHELS, Amsterdam
G. W. RATHENAU, Amsterdam

Sweden

J. O. LINDE, Stockholm
PER-OLOV LÖWDIN, Uppsala

Switzerland

G. BUSCH, Zurich
P. SCHERRER, Zurich

U.S.S.R.

G. S. ZHDANOV, Moscow

PERGAMON PRESS NEW YORK · LONDON · PARIS · LOS ANGELES

NOTES FOR CONTRIBUTORS

I. GENERAL

1. Submission of a paper to the *Journal of the Physics and Chemistry of Solids* will be taken to imply that it represents original research not previously published (except in the form of an abstract or preliminary report), that it is not being considered for publication elsewhere, and that if accepted, it will not be published elsewhere in the same form, in any language, without the consent of the editor-in-chief. It should deal with original research work in the field of the physics and chemistry of solids.

2. Papers should be submitted to the appropriate regional editor (all English-language papers to be sent to the U.S. editor).

3. Papers will be published as quickly as possible after acceptance, and, subject to space being available, should appear in the following issue, if this is due for publication not earlier than three months after the acceptance date. Short communications under the heading of "Letters to the Editor" will receive priority for publication, and will be published in the issue following receipt, if accepted not later than the beginning of the month preceding publication.

4. Fifty free reprints of each paper are supplied. Additional copies can be obtained at a reasonable cost if ordered when proofs are returned. A reprint order form will accompany first proofs.

II. SCRIPT REQUIREMENTS

1. Papers submitted should be concise and written in a readily understandable style. Scripts should be typed and double spaced and submitted in duplicate to facilitate refereeing.

It will be appreciated if authors clearly indicate any special characters used. An abstract, not exceeding 200 words, should be provided in the language of the paper. French and German papers should be submitted with English abstract and titles, but if this is not possible the abstract will be translated by the publishers. To conserve space, authors are requested to mark less important parts of the paper, such as details of experimental technique, methods, mathematical derivations, etc. for printing in small type. The technical description of methods should be given in detail only when such methods are new. Authors will receive proofs for correction when their papers are first set; page proofs will be sent only when the amount of alteration makes it advisable.

2. Illustrations should not be included in the typescript of the paper, and legends should be typed on a separate sheet. Line drawings which require redrawing should include all relevant details and clear instructions for the draughtsman. If figures are already well drawn it may be possible to reproduce them direct from the originals, or from good photo-prints if these can be provided. It is not possible to reproduce from prints with weak lines. Illustrations for reproduction should normally be about twice the final size required. The lettering should be sufficiently large and bold to permit this reduction. Photographs should only be included where they are essential.

3. Tables and figures should be so constructed as to be intelligible without reference to the text. Every table and column should be provided with an explanatory heading. Units of measure must always be clearly indicated. The same data should not be published in both tables and figures. The following standard symbols should be used on line drawings since they are easily available to the printers: ○, ●, +, ×, □, ■, △, ▲, ◇, ◆, ▽, ▼.

4. References are indicated in the text by superior numbers in parentheses, and the full reference should be given in a list at the end of the paper in the following form:

1. HILL R., *J. Iron St. Inst.* **158**, 177 (1948).
2. PEARSON C., *The Extrusion of Metals*, p. 137. Chapman and Hall, London (1944).

Abbreviations of journal titles should follow those given for *Physics Abstracts*. It is particularly requested that authors' initials, and appropriate volume and page numbers, should be given in every case.

Footnotes, as distinct from literature references should be indicated by the following symbols—*, †, ‡, ‡‡, commencing anew on each page; they should *not* be included in the numbered reference system.

5. Due to the international character of the journal no rigid rules concerning notation and spelling will be observed, but each paper should be consistent within itself as to symbols and units.

VOL.
13
1960

VOL
1
19

CONTENTS LISTS FOR VOLUME 13, NOS. 1–4

Volume 13 Numbers 1/2

page

T. P. McLean and R. Loudon: Exciton energy levels in germanium and silicon 1
P. G. de Gennes et J. Villain: Fluctuations d'aimantation et opalescence critique 10
S. Geller, R. M. Bozorth, C. E. Miller and D. D. Davis: Crystal chemical and magnetic studies of garnet systems 28
M. A. Gilleo: Superexchange interaction in ferrimagnetic garnets and spinels which contain randomly incomplete linkages 33
G. A. Alers, J. R. Neighbours and H. Sato: Temperature dependent magnetic contributions to the high field elastic constants of nickel and an Fe–Ni alloy 40
M. Aven and J. A. Parodi: Study of the crytalline transformations in ZnS: Cu, ZnS: Ag and ZnS: Cu, Al 56
R. O. Carlson: Anisotropic diffusion of copper into bismuth telluride 65
A. W. Overhauser: Mechanism of antiferromagnetism in dilute alloys 71
E. F. Apple and J. S. Prener: On the infrared emission in ZnS: Cu—effect of sulfur pressure and aluminum 81
B. S. Gourary and A. A. Maradudin: Absorption and emission spectra of an electron in a one-dimensional deep trap 88
J. E. Hebborn and E. H. Sondheimer: The diamagnetism of conduction electrons in metals 105
R. L. Hansler and W. G. Segelken: Correlation of thermoluminescence in MgO and valence changes of iron and chromium impurities detected by epr 124
R. G. S. Morfee, L. A. K. Staveley, S. T. Walters and D. L. Wigley: A heat capacity study of the movement of the ammonium ion in ammonium stannichloride and stannibromide by comparison of the heat capacities of the ammonium, rubidium and potassium salts 132
Mino Green and K. H. Maxwell: The absorption of oxygen on clean silicon surfaces 145
J. C. Woolley and B. Ray: Solid solution in $A^{II}B^{VI}$ tellurides 151
R. Nitsche and W. J. Merz: Photoconduction in ternary V–VI–VII compounds 154
B. T. Matthias, H. Suhl and E. Corenzwit: Further experiments concerning the spin-lattice interactions in superconductors 156

Letters to the Editor:

P. Flubacher, A. J. Leadbetter and J. A. Morrison: The limiting low-temperature behavior of the heat capacity of graphite 160
R. Dalven: Absorption edge in β-silicon carbide 163
R. E. Michel: The temperature-dependence of electron spin-lattice relaxation times in ruby 165
G. Shirane, S. J. Pickart and R. Newnham: Neutron-diffraction study of Ti_2O_3 167
M. Bernard and B. Leduc: Jonctions de germanium p–n plastiquement deformées 168

Contents lists of Volume 9, Nos. 1–4 171

Volume 13 Numbers 2/3

Y. Sasaki: p–i–n junction in the anodic oxide film of tantalum 177
C. S. Fuller, F. H. Doleiden and Katherine Wolfstirn: Reactions of group III acceptors with oxygen in silicon crystals 187
R. Blinc: On the isotopic effects in the ferroelectric behaviour of crystals with short hydrogen bonds . . 204
J. S. Wollam and W. E. Wallace: Magnetic susceptibility, heat capacity and third-law entropy of $MgNi_2$. 212
G. Leman: Influence des conditions aux limites sur la densité électronique dans un réseau périodique . . 221
M. Ericson et B. Jacrot: Etude expérimentale de la diffusion critique des neutrons dans le fer . . . 235
O. P. Manley: A method of evaluating diffusion coefficients in crystals 244
D. M. Warschauer and D. C. Reynolds: Mechanically-excited emission in cadmium sulfide . . . 251
L. Rimai and N. Bloembergen: Nuclear magnetic resonance in alkali alloy systems, NaK and NaRb . . 257

Foreword for the ten papers which were presented at the Mellon Institute Symposium 271

G. J. Dienes: Introduction—defects in silicas 272
W. Primak: A review of the gross structural effects of energetic atomic particles on vitreous and crystalline silica 279

VOL. 13 1960

page

P. W. LEVY: Reactor and gamma-ray induced coloring of Corning fused silica 287
C. M. NELSON and J. H. CRAWFORD, JR.: Optical absorption in irradiated quartz and fused silica . . 296
G. W. ARNOLD, JR.: Defects in natural and synthetic quartz 306
A. J. COHEN: Substitutional and interstitial aluminum impurity in quartz, structure and color center interrelationships 321
D. L. WOOD: Infrared absorption of defects in quartz 326
C. S. BROWN and L. A. THOMAS: The effect of impurities on the growth of synthetic quartz . . 337
F. AUGUSTINE and D. R. HALE: Topography and etch patterns of synthetic quartz 344
F. M. ERNSBERGER: Structural effects in the chemical reactivity of silica and silicates 347
J. C. KING: The anelasticity of natural and synthetic quartz at low temperatures—abstract 352

Letter to the Editor:

E. DANIEL: Sur la résistivité et le déplacement de Knight d'alliages liquides de sodium 353

Notice 356

R. F. BREBRICK: The minimum molar Gibbs free energy and small deviations from stoichiometry in crystalline binary compounds—Erratum 356

Books received 356

Tome 13 Nos 1/2

T. P. MCLEAN et R. LOUDON: Niveaux d'énergie d'exciton dans le germanium et le silicium 1
P. G. DE GENNES et J. VILLAIN: Fluctuations d'aimantation et opalescence critique 10
S. GELLER, R. M. BOZORTH, C. E. MILLER et D. D. DAVIS: Etudes cristallines chimiques et magnétiques des grenats 28
M. A. GILLEO: Interaction de superéchange dans les grenats et les spinels ferrimagnétiques qui contiennent des liaisons incomplètes réparties au hasard 33
G. A. ALERS, J. R. NEIGHBOURS et H. SATO: Effet de la température sur les contributions magnétiques aux consantes élastiques à grand champ du nickel et d'un alliage Fe–Ni 40
M. AVEN et J. A. PARODI: Etude des transformations cristallines de ZnS: Cu, ZnS: Ag et ZnS: Cu, Al . 56
R. O. CARLSON: Diffusion anisotrope du cuivre dans le tellurure de bismuth 65
A. W. OVERHAUSER: Mécanisme de l'antiferromagnétisme dans les alliages dilués 71
E. F. APPLE et J. S. PRENER: Sur l'émission infrarouge dans ZnS : Cu, influence de la pression de vapeur du soufre et de la teneur en phosphore de l'aluminum 81
B. S. GOURARY et A. A. MARADUDIN: Spectre d'absorption et d'émission pour un électron dans un piège profond à une dimension 88
J. E. HEBBORN et E. H. SONDHEIMER: Diamagnétisme des électrons de conduction dans les métaux . . 105
R. L. HANSLER et W. G. SEGELKEN: Corrélation entre la thermoluminescence dans MgO et le changement de valence des impuretés de fer et de chrome. Etude par résonance paramagnétique électronique . . 124
R. G. S. MORFEE, L. A. K. STAVELY, S.T. WALTERS et D. L. WIGLEY: Etude du mouvement de l'ion ammonium dans le stannichlorure et stannibromure d'ammonium par comparaison des capacitiés calorifiques des sels d'ammonium de, rubidium et de potassium 132
MINO GREEN et K. H. MAXWELL: L'adsorption de l'oxygene par des surfaces propres de silicium . . 145
J. C. WOOLEY et B. RAY: Solutions solides dans les tellurures $A^{II}B^{VI}$ 151
R. NITSCHE et W. J. MERZ: Photoconduction dans les composés ternaires V–VI–VII 154
B. T. MATTHIAS, H. SUHL et E. CORENZWIT: Nouvelles expériences sur les interactions spin-électron dans les supraconducteurs 156

Lettres au Rédacteur:

P. FLUBACHER, A. J. LEADBETTER et J. A. MORRISON: Le comportement limite à basse température de la capacité calorifique du graphite 160
R. DALVEN: Limite d'absorption dans le carbure de silicium β 163
R. E. MICHEL: La dépendance en température des temps de relaxation électroniques spin réseau du rubis 165
G. SHIRANE, S. J. PICKART et R. NEWNHAM: Etude de Ti_2O_3 par la diffraction des neutrons 167
M. BERNARD et B. LEDUC: Jonctions de germanium p–n plastiquement deformées 168

Tables des Matières du Volume 9, Nos. 1–4 171

Tome 13 Nos 3/4

page

Y. SASAKI: Jonction *p–i–n* dans un film anodique d'oxyde de tantale 177
C. S. FULLER, F. H. DOLEIDEN et KATHERINE WOLFSTIRN: Réaction des accepteurs du groupe III avec l'oxygène dans les cristaux de silice 187
R. BLINC: Les effets isotopiques dans le comportement ferroélectrique des cristaux à liaisons hydrogéne courtes 204
J. S. WOLLAM et W. E. WALLACE: Susceptibilité magnétique, capacité calorifique et entropie de zéro de $MgNi_2$ 212
G. LEMAN: Influence des conditions aux limites sur la densité électronique dans un réseau périodique . . 221
M. ERICSON et B. JACROT: Étude expérimentale de la diffusion critique des neutrons dans le fer . . . 235
O. P. MANLEY: Une méthode d'évaluation des coefficients de diffusion des cristaux 244
D. M. WARSCHAUER et D. C. REYNOLDS: Emission excitée mécaniquement dans le sulfure de cadmium . 251
L. RIMAI et N. BLOEMBERGEN: Résonance magnétique nucléaire dans les alliages alcalins NaK et NaRb . 257

Préface aux dix rapports présentés au Symposium de l'Institut Mellon 271

G. J. DIENES: Introduction. Défauts dans les silices 272
W. PRIMAK: Une revue des effets macroscopiques de particules atomiques énergiques sur la structure de la silice vitreuse et cristalline 279
P. W. LEVY: Coloration de la silice fondue induite par les rayons gamma et le rayonnement de réacteurs . 287
C. M. NELSON et J. H. CRAWFORD, JR.: Absorption optique dans le quartz irradié et la silice fondue . . 296
G. W. ARNOLD, JR.: Défauts dans le quartz naturel ou synthétique 306
A. J. COHEN: L'aluminium interstitiel et substitutionnel dans le quartz. Les relations entre la structure et les centres colorés 321
D. L. WOOD: Absorption infrarouge par les défauts dans le quartz 326
C. S. BROWN et L. A. THOMAS: Les effets des impuretés sur a croissance du quartz synthétique . . 337
T. AUGUSTINE et D. R. HALE: Etude de la surface et des figures d'attaque du quartz synthétique . . 344
F. M. ERNSBERGER: Effets de structure dans la réactivité chimique de la silice et des silicates . . 347
J. C. KING: Anélasticité du quartz naturel et synthétique aux basses temperatures—Abrègé . . . 352

Lettre au Rédacteur:
E. DANIEL: Sur la resistivité et le déplacement de Knight d'alliages liquides de sodium 353

Note 356

R. F. BREBRICK: Le minimum d'enérgie molaire libre de Gibbs et de petits déviations de stoichiométrie dans les composés cristallins—Erratum 356

Ouvrages reçus 356

Band 13 Nr. 1/2

T. P. MCLEAN und R. LOUDON: Excitonen-Niveaus in Germanium und Silizium 1
P. G. DE GENNES und J. VILLAIN: Die Magnetisierungsschwankungen und das kritische Opalisieren . . 10
S. GELLER, R. M. BOZORTH, C. E. MILLER und D. D. DAVIS: Kristall-chemische und magnetische Studien der Granatsysteme 28
M. A. GILLEO: Extra-Austauschwechselwirkung in eisenmagnetischen Granaten und Spinellen, welche eine willkürlich unvollkommene magnetische Verbindung besitzen 33
G. A. ALERS, J. R. NEIGHBOURS und H. SATO: Die von der Temperatur abhängigen magnetischen Beiträge zu den hohen elastischen Feldkonstanten in Nickel und Fe–Ni-Legierungen 40
M. AVEN und J. A. PARODI: Studium der Umkristallisation in ZnS: Cu, ZnS: Ag und ZnS: Cu, Al . . 56
R. O. CARLSON: Die anisotropische Diffusion von Kupfer in Wismuttellurid 65
A. W. OVERHAUSER: Der antiferromagnetische Vorgang in flüssigen Legierungen 71
E. F. APPLE und J. S. PRENER: Über die Ultrarotemission des ZnS : Cu-Effekts bei Schwefeldruck und Aluminium 81
B. S. GOURARY und A. A. MARADUDIN: Absorptions und Emissions-spektren eines tiefen eindimensionalen Elektroneneinfangs 88
J. E. HEBBORN und E. H. SONDHEIMER: Der Diamagnetismus von Leitelektronen in Metallen 105
R. L. HANSLER und W. G. SEGELKEN: Die Beziehung zwischen der Thermolumineszenz in MgO und des Valenzwechsels von Eisen- und Chromunreinigkeiten, welche durch paramagnetische Elektronenresonanz festgestellt werden können 124

VOL. 13 1960

page

R. G. S. Morfee, L. A. K. Stavely, S. T. Walters und D. L. Wigley: Studium der Wärmekapazität der Bewegung von Ammoniumionen in Zinnammoniumchlorid und Zinnammoniumbromid durch Vergleichung der Wärmekapazitäten von Ammonium-, Rubidium und Kaliumsalzen 132
Mino Green und K. H. Maxwell: Die Sauerstoffabsorption auf reinen Siliziumoberflächen . . . 145
J. C. Wooley und B. Ray: Die feste Lösung von $A^{II}B^{VI}$ Telluriden 151
R. Nitsche und W. J. Merz: Lichtelektrische Leitfähigkeit von V–VI–VII-Dreifachverbindungen . . 154
B. T. Matthias, H. Suhl und E. Corenzwit: Weitere Versuche über die Wechselwirkung von Spinelektronen in Supraleitern 156

Briefe an die Herausgeber:

P. Flubacher, A. J. Leadbetter und J. A. Morrison: Die begrenzende Eigenschaft der Tieftemperatur der Wärmekapazität des Graphits 160
R. Dalven: Die Absorptionskante des β-Siliziumkohlenstoffes 163
R. E. Michel: Die Temperaturabhängigkeit der Relaxationszeiten des Elektronenspingitters in Rubine . 165
G. Shirane, S. J. Pickart und R. Newnham: Studium der Neutronendiffraktion in Ti_2O_3 167
M. Bernard und B. Leduc: Plastisch deformierte *p–n*-Verbindungen des Germaniums 168
Inhaltsverzeichnis des 9. Bandes, Nr. 1–4 171

Band 13 Nr. 3/4

Y. Sasaki: Die *p–i–n*-Verbindung in anodischen Tantaloxydschichten 177
C. S. Fuller, F. H. Doleiden und Katherine Wolfstirn: Reaktionen der 111 Akzeptorgruppe mit Sauerstoff in Siliziumkristallen 187
R. Blinc: Über die isotopischen Wirkungen der ferroelektrischen Eigenschaft von Kristallen mit kurzen Wasserstoffbindungen 204
J. S. Wollam und W. E. Wallace: Die magnetische Aufnahmefähigkeit, die Wärmekapazität und die Entropie von $MgNi_2$ des dritten Gesetzes 212
G. Leman: Die Wirkung des Zustandes an den Grenzen auf die Elektronendichte in einen periodischen Netz 221
M. Ericson und B. Jacrot: Experimentelles Studium der kritischen Neutronendiffusion im Eisen . . 235
O. P. Manley: Eine Methode zur Abschätzung der Diffusions-koeffizienten in Kristallen . . . 244
D. M. Warschauer und D. C. Reynolds: Mechanisch angeregte Emission in Kadmiumsulfid . . . 251
L. Rimai und N. Bloembergen: Die kernmagnetische Resonanz in den alkalischen Legierungssystemen von NaK und NaRb 257

Vorwort zu den zehn Abhandlungen, welche zur Tagung am Mellon Institut eingeliefert worden waren . 271

G. J. Dienes: Einführung—Mängel in der Kieselerde 272
W. Primak: Eine Übersicht über die Grossstrukturwirkung der energiereichen Atomteilchen auf glasartige und kristallartige Kieselerden 279
P. W. Levy: Die durch Reaktion und durch Gammastrahlen induzierte Färbung von geschmolzener Kieselerde nach der Corning Methode 287
C. M. Nelson und J. H. Crawford, Jr.: Die optische Absorption in bestrahltem Quarz und in geschmolzener Kieselerde 296
G. W. Arnold, Jr.: Mängel im natürlichen und im synthetischen Quarz 306
A. J. Cohen: Substitutions- und Zwischengitteraluminiumreinheiten in Quarz; Struktur- und Farbzentrumsbeziehungen 321
D. L. Wood: Ultrarotabsorption der Mängel im Quarz 326
C. S. Brown und L. A. Thomas: Die Wirkung von Unreinheiten auf das Wachstum des synthetischen Quarzes 337
F. Augustine und D. R. Hale: Oberflächengestaltung und Ätzfiguren des synthetischen Quarzes . . 344
F. M. Ernsberger: Aufbauwirkung der chemischen Reaktionsfähigkeit in der Kieselerde und in Silikaten 347
J. C. King: Die Anelastizität des natürlichen und synthetischen Quarzes bei Tieftemperaturen—Auszug . 352

Brief an die Herausgeber:

E. Daniel: Über den spezifischen Widerstand und über die Verschiebung der flüssigen Natriumlegierungen nach der Methode Knight 353

Zur Beachtung 356

R. F. Brebrick: Die minimale freie Molekularenergie nach Gibbs und die geringen Abweichungen von der Stöchiometrie in kristallartigen Zweistoffverbindungen—Druckfehler 356

Bücher erhalten 356

Том 13 На. 1/2

page

Т. П. МакЛин и Р. Лоудон: Экситонные уровни энергии в германии и кремнии . . . 1
П. де Геннес и Ж. Виллен: Флуктуации намагниченности и критическая опалесценция . . 10
С. Геллер, Р. М. Бозорт, К. Миллер и Д. Д. Девис: Кристаллохимическое и магнитное исследование гранатовых систем $\{YCa_2\}[M_2^{4+}](Fe_3)O_{12} - \{Y_3\}[Fe_2](Fe_3)O_{12}$, M = Zr, Hf. . . 28
М. А. Гиллео: Суперобменное взаимодействие в ферримагнитных гранатах и шпинелях, содержащих произвольно неполные связи 33
Г. А. Алерс, Дж. Нейборс и Г. Сато: Зависящие от температуры магнитные вклады в упругие константы никкеля и сплава Fe–Ni в сильных полях 40
М. Авен и Дж. Пароди: Изучение кристаллических превращений в ZnS:Cu, ZnS:Ag и ZnS:Cu, Al. 56
Р. О. Карлсон: Анизотропная диффузия меди в теллурид висмута 65
А. В. Оверхаузер: Механизм антиферромагнетизма разведенных растворов . . . 71
Е. Ф. Эппль и Дж. Пренер: Об инфракрасной эмиссии в ZnS:Cu — влияние давления серы и алюминия 81
Б. С. Гурари и А. А. Марадудин: Спектры поглошения и испускания электрона в глубокой одномерной ловушке 88
Дж. Е. Гебборн и Е. Г. Зондгеймер: Диамагнетизм кондукционных электронов в металлах . 105
Р. Л. Ганслер и В. Зегелькен: Связь между термолюминесценцией MgO и изменениями валентности примесей железа и хрома, обнаруживаемых электронным парамагнитным резонансом 124
Р. Морфи, Л. Стевли, С. Вальтерс и Д. Виглей: Термическое исследование движения иона аммония в $(NH_4)_2SnCl_6$ и $(NH_4)_2SnBr_6$ путем сравнения теплоемкостей аммониевых, рубидиевых и калиевых солей 132
М. Н. Грин и К. Г. Максвелл: Адсорбция кислорода на чистой поверхности кремния . 145
Дж. Вуллей и Б. Рей: Твердые растворы теллуридов $A^{II}B^{VI}$ 151
Р. Нитше и В. Мерц: Фотопроводимость в тройных соединениях типа У–У1–УП . . . 154
Б. Т. Маттиас, Г. Сул и Е. Коренцвит: Дальнейшие опыты по спинэлектронному взаимодействию в сверхпроводниках 156

Письма в Редакцию:

П. Флубахер, А. Ледбеттер и Дж. Моррисон: Предельное низко-температурное поведение теплоемкости графита 160
Р. Дальвен: Край поглощения в β-карбиде кремния 163
Р. Мишель: Температурная зависимость электронного спин-решеточного времени релаксации в рубине 165
Г. Ширейн, С. Пиккарт и Р. Ньюгем: Нейтроннодиффракционное исследование Ti_2O_3 . 167
М. Бернар и Б. Ледюк: Переходы в пластически деформированном p–n германии . . 168

Оглавления тома 9 выпусков 1–4 171

Том 13 На. 3/4

И. Сазаки: Переход p–i–n в анодной оксидной пленке тантала 177
К. Фуллер, Ф. Долейден и К. Вольфштирн: Реакции акцепторов третьей группы с кислородом в кристаллах кремния 187
Р. Блинк: Об изотопических эффектах в ферроэлектрическом поведении кристаллов с короткими водородными связями 204
Дж. С. Воллам и В. Е. Уоллес: Магнитная восприимчивость, теплоемкость и энтропия $MgNi_2$ 212
Г. Леман: Влияние граничных условий на электронную плотность в периодической решетке 221
М. Эриксон и Б. Жакро: Экспериментальное исследование критической диффузии нейтронов в железе 235
О. П. Манлей: Метод для определения коэффициента диффузии в кристаллах 244
Д. М. Варшауер и Д. Рейнольдс: Механически возбужденная эмиссия сульфида кадмия . 251
Л. Римаи и Б. Блумберген: Ядерный магнитный резонанс в сплавах щелочных металлов NaK и NaRb 257

VOL. 13 1960

page

Предисловие к десяти статьям, представленным на симпозиуме Меллоновского Института . 271

Г. Динес: Введение — дефекты в силикатах 272
В. Примак: Обзор по вопросу о структурных эффектах энергичных атомных частиц в стеклообразных и кристаллических силикатах 279
Н. В. Леви: Окраска плавленных силикатов, индуцированная реакторным или гамма-облучением 287
К. Нельсон, Дж. Кроуфорд, мл.: Оптическое поглощение в облученных кварце и плавленых силикатах 296
Г. В. Арнольд мл.: Дефекты в естественном и синтетическом кварце 306
А. Коэн: Примеси алюминия типа внедрения и типа замещения в кварце; связь между структурой и центрами окраски 321
Д. Л. Вуд: Инфракрасное поглощение на дефектах в кварце 326
К. Броун и Л. А. Томас: Влияние примесей на рост синтетического кварца 337
Ф. Огюстен и Д. Р. Хейл: Топография и фигуры травления синтетического кварца . . 344
Ф. М. Эрнсбергер: Структурные эффекты в химической реактивности кремния и силикатов 347
Дж. С. Кинг: Неупругость природного и синтетического кварца при низких температурах — Абстракт 352

Письмо в Редакцию:
Е. Даниель: О сопротивлении и смещении Найта в жидких сплавах натрия . . . 353

Объявление 356

Р. Ф. Бребрик: Минимальная молярная свободная энергия "Гиббс" и небольшие отклонения от стехиометрии в кристаллических соединениях — Опечатка 356

Полученные книги 356

AUTHOR INDEX TO VOLUME 13

Alers, G. A. 40
Apple, E. F. 81
Arnold, G. W., Jr. 306
Augustine, F. 344
Aven, M. 56
Bernard, M. 168
Blinc, R. 204
Bloembergen, N. 257
Bozorth, R. M. 28
Brebrick, R. F. 356
Brown, C. S. 337
Carlson, R. O. 65
Cohen, A. J. 321
Corenzwit, E. 156
Crawford, J. H., Jr. 296
Dalven, R. 163
Daniel, E. 353
Davis, D. D. 28
Dienes, G. J. 272
Doleiden, F. H. 187
Ericson, M. 235
Ernsberger, F. M. 347
Flubacher, P. 160
Fuller, C. S. 187
Geller, S. 28
De Gennes, P. G. 10
Gilleo, M. A. 33
Gourary, B. S. 88
Green, M. 145
Hale, D. R. 344
Hansler, R. L. 124
Hebborn, J. E. 105
Jacrot, B. 235
King, J. C. 352
Leadbetter, A. J. 160
Leduc, B. 168
Leman, G. 221
Levy, P. W. 287
Loudon, R. 1
Manley, O. P. 244
Maradudin, A. A. 88
Matthias, B. T. 156
Maxwell, K. H. 145
Mclean, T. P. 1
Merz, W. J. 154
Michel, R. E. 165
Miller, C. E. 28
Morfee, R. G. S. 132
Morrison, J. A. 160
Neighbours, J. R. 40
Nelson, C. M. 296
Newnham, R. 167
Nitsche, R. 154
Overhauser, A. W. 71
Parodi, J. A. 56
Pickart, S. J. 167
Prener, J. S. 81
Primak, W. 279
Ray, B. 151
Reynolds, D. C. 251
Rimai, L. 257
Sasaki, Y. 177
Sato, H. 40
Segelken, W. G. 124
Shirane, G. 167
Sondheimer, E. H. 105
Staveley, L. A. K. 132
Suhl, H. 156
Thomas, L. A. 337
Villain, J. 10
Wallace, W. E. 212
Walters, S. T. 132
Warschauer, D. M. 251
Wigley, D. L. 132
Wolfstirn, K. 187
Wollam, J. S. 212
Woolley, J. C. 151
Wood, D. L. 326

VOL.
13
1960

VO
1
19

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 13. pp. 1–9 Printed in Great Britain.

EXCITON ENERGY LEVELS IN GERMANIUM AND SILICON

T. P. McLEAN

Royal Radar Establishment, Malvern, England

and

R. LOUDON

Clarendon Laboratory, Oxford, England

(*Received* 17 *July* 1959)

VOL. 13 1960

Abstract—The energies of the lowest-lying levels of the direct exciton in germanium and the indirect excitons in germanium and silicon have been calculated in the effective-mass approximation by a variational procedure. These energies are compared with the experimental values obtained from optical absorption experiments.

1. INTRODUCTION

EFFECTS attributable to the formation of excitons have recently been observed by a group at R.R.E.[1] in the optical absorption spectra of both silicon and germanium and by a group at Lincoln Laboratory[2] in the spectrum of germanium. These observations allow values of the binding energy of the lowest-lying exciton states to be determined. The values found are in rough agreement with estimates made by DRESSELHAUS[3] on the basis of an effective-mass approximation for the exciton.

The description of exciton energy levels by means of an effective-mass approximation was first proposed by WANNIER[4] and developed by him for the case of simple valence and conduction bands. DRESSELHAUS[3] has reformulated this treatment and extended it to cover the cases in which the band extrema are degenerate and occur at arbitrary points in the Brillouin zone. It follows from this work that the degeneracy of the valence-band edge in germanium and silicon leads to a description of the exciton system in these substances by a set of coupled differential equations which are very similar to the equations describing acceptor impurities in germanium and silicon.[5] These equations have recently been solved by a variational method.[6, 7] We present in this paper the details and results of a similar calculation of the exciton energy levels. The direct exciton states have a degeneracy similar to that found in the acceptor states.[5] This degeneracy is partially lifted for the indirect excitons as the electrons bound in these states are associated with several equivalent conduction-band minima with anisotropic mass. A detailed group theoretical discussion of these excitons is given here by a method which is general and could be used in more complicated cases.

In germanium two types of exciton have been observed, one type being produced by direct optical transitions and the other by indirect optical transitions involving the creation or annihilation of a phonon. We shall refer to these as direct and indirect excitons respectively. In silicon only indirect excitons have been observed. We shall limit our considerations to these three cases.

2. THE EXCITON HAMILTONIAN

The exciton consists of a hole in the valence band close to $\boldsymbol{k} = 0$ bound to an electron close to a conduction-band minimum at wave-vector $\boldsymbol{k} = \boldsymbol{k}_c$. In both germanium and silicon the valence band is threefold degenerate at $\boldsymbol{k} = 0$ when spin is ignored, and when spin is taken into account the sixfold degenerate band is split by spin-orbit interaction into a fourfold and a twofold band with

the fourfold band uppermost. The conduction bands have only spin degeneracy, and since we shall not take into account the spin-orbit interaction between the electron and hole, we may ignore the conduction-band spin and take the bands to be non-degenerate.

The wave-function of an exciton at the exciton-band minimum may be written:

$$\Psi_{n\mathbf{k}_c}(\mathbf{r}_e,\mathbf{r}_h) = \sum_i F^i_{n\mathbf{k}_c}(\mathbf{r})\psi^i_{v0}(\mathbf{r}_h)\psi_{c\mathbf{k}_c}(\mathbf{r}_e) \quad (2.1)$$

where $\psi_{c\mathbf{k}_c}(\mathbf{r}_e)$ and $\psi^i_{v0}(\mathbf{r}_h)$ are the Bloch functions for the conduction- and valence-band edges, $\mathbf{r}_e$ is the position of the electron, $\mathbf{r}_h$ that of the hole and:

$$\mathbf{r} = \mathbf{r}_e - \mathbf{r}_h. \quad (2.2)$$

The i summation runs over the degenerate valence-band states at $\mathbf{k} = 0$, and the subscript n denotes the internal state of the exciton. The functions $F^i_{n\mathbf{k}_c}(\mathbf{r})$ describe the relative motion of the electron and hole.

We are going to classify the exciton states group theoretically, so we require to know the group of transformations under which the exciton Hamiltonian is invariant. The Hamiltonian consists of a sum of three terms representing the electron–hole interaction energy and the effective-mass kinetic energies of a hole close to $\mathbf{k} = 0$ in the valence band and an electron close to $\mathbf{k} = \mathbf{k}_c$ in the conduction band. The presence of the kinetic-energy terms in the exciton Hamiltonian requires that the transformations applied to $\mathbf{r}_e$ and $\mathbf{r}_h$ must be contained in the symmetry groups associated with the points $\mathbf{k} = \mathbf{k}_c$ and $\mathbf{k} = 0$ respectively. Further, the interaction energy depends on the electron–hole separation $|\mathbf{r}|$, and this term is maintained invariant only if the transformations simultaneously applied to the electron and hole co-ordinates $\mathbf{r}_e$ and $\mathbf{r}_h$ are identical. We therefore have the result that the group of the Hamiltonian includes those operations in which a transformation contained in the intersection of the groups of the $\mathbf{k}$-vectors of the valence-band maximum and the conduction-band minimum is simultaneously applied to the electron and hole co-ordinates. For the cases we shall consider, the group of $\mathbf{k}_c$, G say, is either equal to that of $\mathbf{k} = 0$ or is a subgroup of it, so that the required intersection is just the group G itself.

Now consider the transformation properties of the exciton wave-function (2.1). Under the operations of the group of the Hamiltonian, it is clear that the $\psi^i_{v0}(\mathbf{r}_h)$ and $\psi_{c\mathbf{k}_c}(\mathbf{r}_e)$ are basis functions for representations of the group G, the former functions generating a representation corresponding to the point $\mathbf{k} = 0$ and the latter a representation associated with the point $\mathbf{k} = \mathbf{k}_c$. It remains to consider the effect of these operations on the functions $F^i_{n\mathbf{k}_c}(\mathbf{r})$. These functions satisfy a set of coupled differential equations derived by DRESSELHAUS[3], which are defined as follows for the case of germanium and silicon. The valence bands are described for small values of $\mathbf{k}$ by the well-known 6×6 matrix given, e.g., by KOHN (Ref. 5, equation (8.4)) and which we denote by $H^v(\mathbf{k})$, while the conduction-band energy is represented by $H^c(\mathbf{k}-\mathbf{k}_c)$. We take the interaction between the electron and hole to be of the Coulomb type modified by the static dielectric constant κ, i.e. $-e^2/\kappa r$. Then the Hamiltonian for the relative motion of the electron and hole is:

$$H^{Ex}_{ij} = \left[H^c(-i\nabla) - \frac{e^2}{\kappa r}\right]\delta_{ij} - H^v_{ij}(-i\nabla) \quad (2.3)$$

where ∇ is the gradient operator in $\mathbf{r}$-space, and the $F^i_{n\mathbf{k}_c}(\mathbf{r})$ satisfy the set of equations:

$$\sum_j H^{Ex}_{ij}F^j_{n\mathbf{k}_c}(\mathbf{r}) = E_{n\mathbf{k}_c}F^i_{n\mathbf{k}_c}(\mathbf{r}). \quad (2.4)$$

It is clear from the definition (2.2) that under the operations of the group of the exciton Hamiltonian $\mathbf{r}$ experiences only the rotational parts of the operators of G, since the same translation applied to $\mathbf{r}_e$ and $\mathbf{r}_h$ cancels when the difference is taken to obtain the transformed value of $\mathbf{r}$. The Hamiltonian (2.3) for the relative motion of the electron and hole is therefore invariant under the operations of the point-group, G_o say, isomorphic with the space-group G, these operations being applied now to the co-ordinate $\mathbf{r}$. Hence, under the operations of the group of the total exciton Hamiltonian, the wave-functions $F^i_{n\mathbf{k}_c}(\mathbf{r})$ for the relative motion are basis functions for representations of the point group G_o. However, space-group irreducible representations corresponding to the point $\mathbf{k} = 0$ can be formed from those of the isomorphic point-group by associating a factor unity with the translational part of the space-group operation whose representation is required.[8] The

functions $F^i_{n\boldsymbol{k}_c}(\boldsymbol{r})$ therefore form a basis for a $\boldsymbol{k}=0$ representation of the space-group G, where it is important to remember that the transformations of this group are applied not to the co-ordinate $\boldsymbol{r}$ but to the co-ordinates $\boldsymbol{r}_e$ and $\boldsymbol{r}_h$.

We have shown that under the operations of the group of the exciton Hamiltonian the three components of the exciton wave-function (2.1) form bases for representations of the group G of the wave-vector of the conduction-band minimum, the first two components generating representations corresponding to the point $\boldsymbol{k}=0$, and the third component giving a representation associated with the point $\boldsymbol{k}=\boldsymbol{k}_c$. The transformation properties of the total exciton wave-function $\Psi_{n\boldsymbol{k}_c}(\boldsymbol{r}_e, \boldsymbol{r}_h)$ are determined by calculating the direct product of the three representations.

We now consider the form of the Hamiltonian for the relative motion of the electron and hole for each of the three cases in which we are interested.

(a) *The direct exciton in germanium*

For the case of germanium the splitting between the fourfold and the twofold degenerate valence bands is about 0·3 eV,[9] which is about 300 times the observed exciton binding energy. We may therefore neglect the contribution to the exciton wave-function from the twofold degenerate band. Hence we take just the upper left-hand 4×4 part of the matrix $H^v(\boldsymbol{k})$, which describes the full valence-band structure in germanium for small values of $\boldsymbol{k}$, and denote this 4×4 matrix by $H^v_4(\boldsymbol{k})$.

The conduction band is spherical for small values of $\boldsymbol{k}$ and its energy is:

$$H^c(\boldsymbol{k}) = E_o + \frac{\hbar^2k^2}{2m^*} \qquad (2.5)$$

where $m^*=0{\cdot}037m$ is the effective mass, and is known from magneto-absorption measurements,[10] m being the free electron mass. We are measuring energies from the valence-band edge, so that E_o is the energy gap at $\boldsymbol{k}=0$, i.e. the direct energy gap. Substitution of these expressions for $H^v(\boldsymbol{k})$ and $H^c(\boldsymbol{k})$ into (2.3) gives the Hamiltonian for the relative motion of the direct exciton in germanium. The binding energy of the exciton in state n is E_o-E_{no}.

The Bloch functions for the four degenerate valence bands at $\boldsymbol{k}=0$, i.e. $\psi^i_{vo}(\boldsymbol{r})$ $i=1, 2, 3, 4$, form a basis for the irreducible representation Γ^+ of the full cubic double group,[11] while the conduction-band Bloch function $\psi_{co}(\boldsymbol{r})$ transforms as Γ'_2 neglecting spin.[12] We use the notation for irreducible representations given by BOUCKAERT *et al.*[13] and ELLIOTT[14], and we shall also make use of the tables in these two articles. We note that the product of Bloch functions in (2.1) transforms like $\Gamma^+_8\times\Gamma'_2=\Gamma^-_8$, which is a four-dimensional irreducible representation. To obtain the representation by which the total exciton wave-function transforms, Γ^-_8 must be multiplied by the appropriate representation for $F_{no}(\boldsymbol{r})$, whose nature we consider in Section 3.

(b) *The indirect exciton in germanium*

The conduction band in germanium has minima at the zone edges in the (111) and similar directions. The indirect exciton consists of an electron in the conduction band at one of these minima and a hole in the valence band at $\boldsymbol{k}=0$. The constant-energy surfaces close to the (111) direction minimum are ellipsoids of revolution about the (111) direction, and it is convenient to take a Cartesian co-ordinate system (1, 2, 3) in which the 3-axis points along this direction. For $\boldsymbol{k}$ close to $\boldsymbol{k}_c$, the conduction-band energy is then:

$$H^c(\boldsymbol{k}-\boldsymbol{k}_c) = E_G+\frac{\hbar^2}{2m_t}(\Delta k_1^2+\Delta k_2^2)+\frac{\hbar^2}{2m_l}\Delta k_3^2 \qquad (2.6)$$

where $m_l=1{\cdot}60m$ and $m_t=0{\cdot}0813m$ are known from cyclotron-resonance measurements and $\Delta\boldsymbol{k}=\boldsymbol{k}-\boldsymbol{k}_c$; E_G is the minimum energy gap between the valence and conduction bands—the indirect energy gap.

For $H^v(\boldsymbol{k})$ we can again neglect the contribution to the exciton wave-function from the twofold degenerate valence bands and take just the upper left-hand 4×4 block of the complete valence-band matrix. The co-ordinate axes (x, y, z) for $H^v_4(\boldsymbol{k})$ are normally taken in the crystal directions (100), (010) and (001). For the indirect exciton Hamiltonian it is more convenient to apply to $H^v_4(\boldsymbol{k})$ a unitary transformation which alters the reference axes to (1, 2, 3). We denote this transformed matrix by $\tilde{H}^v_4(\boldsymbol{k})$. Substitution of the above valence- and conduction-band Hamiltonians into (2.3) gives the exciton Hamiltonian. In state n the exciton binding energy is $E_G-E_{n\boldsymbol{k}_c}$, where the $E_{n\boldsymbol{k}_c}$ are the eigenvalues of this Hamiltonian. The energy levels of

the excitons with wave-vectors corresponding to the other conduction-band minima are degenerate with the $E_{n\boldsymbol{k}_c}$.

The Bloch functions at $\boldsymbol{k}_c$ form bases for the irreducible representations of the symmetry group L, which is a subgroup of the full cubic group Γ. The L representations corresponding to $\boldsymbol{k}=0$ and $\boldsymbol{k}=\boldsymbol{k}_c$ are different, and we must therefore distinguish between them by quoting the $\boldsymbol{k}$-vector. The compatibility relation between Γ_8^+ and L representations is:

$$\Gamma_8^+ \rightarrow L_4^+(0)+L_5^+(0)+L_6^+(0).$$

$L_6^+(0)$ is two-dimensional and $L_4^+(0)$ and $L_5^+(0)$ are both one-dimensional, being degenerate by time reversal. The conduction-band Bloch function at $\boldsymbol{k}_c$ transforms as the scalar representation $L_1(\boldsymbol{k}_c)$, so that the product of Bloch functions in (2.1) transforms either as $L_4^+(\boldsymbol{k}_c)+L_5^+(\boldsymbol{k}_c)$ or as $L_6^+(\boldsymbol{k}_c)$. Apart from accidental degeneracy, the exciton states in this case cannot have a degeneracy greater than twofold.

(c) *The indirect exciton in silicon*

In silicon the experimentally determined binding energy of the indirect exciton is about 0·01 eV, which is about one-third of the estimated spin-orbit splitting of the valence bands at $\boldsymbol{k}=0$. The twofold degenerate valence bands may therefore make a significant contribution to the exciton state. On account of this one should use the complete 6×6 valence-band matrix $H^v(\boldsymbol{k})$ in forming the exciton Hamiltonian. However, this leads to rather lengthy computations, and we have therefore carried out calculations in two extreme limits, firstly proceeding as in the case of germanium and considering only contributions from the top four valence-band states (infinite spin-orbit splitting), and secondly taking the spin-orbit splitting to be zero, which corresponds to neglecting the valence-band spin.

The conduction-band minima in silicon lie in the (100) and similar directions inside the first zone. Near the minimum in the (001) direction situated at $\boldsymbol{k}=\boldsymbol{k}_c$, the surfaces of constant energy are ellipsoids of revolution about the (001) direction, the energy being given by:

$$H^c(\boldsymbol{k}-\boldsymbol{k}_c) = E_G+\frac{\hbar^2}{2m_t}(\Delta k_x^2+\Delta k_y^2)+\frac{\hbar^2}{2m_l}\Delta k_z^2 \quad (2.7)$$

where $m_l=0{\cdot}98m$ and $m_t=0{\cdot}19m$ are known from cyclotron-resonance measurements, and the co-ordinate axes (x, y, z) are defined in Section 2(*b*).

Assuming infinite spin-orbit splitting, we take for the valence-band matrix in (2.3) the matrix $H_4^v(\boldsymbol{k})$ defined above. The Bloch functions at $\boldsymbol{k}_c$ now form bases for irreducible representations of the Δ subgroup of the full cubic group, the conduction-band Bloch function transforming as $\Delta_1(\boldsymbol{k}_c)$. The Δ representations corresponding to $\boldsymbol{k}=0$ and $\boldsymbol{k}=\boldsymbol{k}_c$ are not identical, so we display the $\boldsymbol{k}$-vectors of the representations in parenthesis. The compatibility relation between Γ_8^+ and Δ group representations is:

$$\Gamma^+ \rightarrow \Delta_6(0)+\Delta_7(0).$$

If we assume zero spin-orbit splitting, we must take for the valence-band matrix in (2.3) the 3×3 matrix which describes the valence-band edge in the absence of spin-orbit coupling[15] and which we shall denote by $H_2^v(\boldsymbol{k})$. In this case the three degenerate valence-band Bloch functions at $\boldsymbol{k}=0$ form a basis for the full cubic group representation Γ_{25}'. When the symmetry group is restricted to Δ, this representation decomposes as follows:

$$\Gamma_{25}' \rightarrow \Delta_2'(0)+\Delta_5(0).$$

Since the conduction-band Bloch function transforms as the scalar representation, the symmetry character of the product of Bloch functions in (2.1) is the same as that of the valence-band Bloch functions alone, except that the $\boldsymbol{k}$-vector of the representation is changed from zero to $\boldsymbol{k}_c$. The exciton states again cannot have degeneracy greater than twofold.

3. THE VARIATIONAL PROCEDURE AND TRIAL FUNCTIONS

In none of the three cases is it possible to solve equations (2.4) exactly for the $F_{n\boldsymbol{k}_c}^i(\boldsymbol{r})$, and we therefore adopt a variational procedure taking trial functions for the F^i. Much labour can be avoided by a suitable choice of trial functions, and we construct our trial functions by using the method of SCHECHTER[7]. We write the $F_{n\boldsymbol{k}_c}^i(\boldsymbol{r})$ in a column matrix which we note by $F_{n\boldsymbol{k}_c}(\boldsymbol{r})$, and we take as trial function for this matrix a linear sum of column matrices, the coefficients in the sum being

VOL. 13 1960

variable parameters. This is the Ritz–Rayleigh variational procedure. We are most interested in the ground state of the exciton for each symmetry type, and we expect that the functions $F_{n\mathbf{k}_c}(\mathbf{r})$ will be dominated by a term of s-like symmetry for such states. However, the form of the exciton Hamiltonian suggests that there should be some admixture of higher angular-momentum states. For example, without these higher states one effectively neglects the fluted nature of the valence bands. Now the exciton Hamiltonians are even under inversion, so that they do not mix states of different parity. Hence the functions included in the ground-state trial function must all be even under inversion. After the s-term we expect that terms of d-like symmetry will be the most important, and in fact we shall not include in the sum matrices of higher angular momentum than d. Taking only the s-term in the trial function, the $F_{n\mathbf{k}_c}(\mathbf{r})$ will transform like scalars, and hence the total exciton wave-function as given by (2.1) will transform in the same way as the product of valence- and conduction-band Bloch functions. For example, in the case of the direct exciton in germanium $\Psi_{no}(\mathbf{r}_e, \mathbf{r}_h)$, where n refers to the ground state, must transform like Γ_8^-. When matrices of d-like symmetry are included in the trial function, these matrices must be such that $\Psi_{no}(\mathbf{r}_e, \mathbf{r}_h)$ still transforms as Γ_8^-. This fact enables us to restrict the number of d-terms which need be included in the trial function. The number of such d-terms may be determined as follows, still taking the direct exciton in germanium as an example. Under the operations of the cubic point-group, the five angular-momentum d-functions form a basis for the point-group irreducible representations corresponding to Γ_{12} and Γ'_{25}. Hence, using terms of d-symmetry in the trial function, the $F^i_{no}(\mathbf{r})$ form a basis for the irreducible representations Γ_{12} and Γ'_{25} of the full cubic space-group. We emphasize again that the transformations of this group are to be applied simultaneously to $\mathbf{r}_e$ and $\mathbf{r}_h$, not to $\mathbf{r}$. It follows that the products of three functions appearing in the exciton wave-function (2.1) form bases for the following representations:

$$(\Gamma_{12}+\Gamma'_{25})\times\Gamma_8^- = 2(\Gamma_6^-+\Gamma_7^-)+3\Gamma_8^-.$$

The representation Γ_8^- occurs three times in the reduction of the direct product. This shows that there are three different matrices of d-symmetry which can be included in the trial function for $F_{no}(\mathbf{r})$ such that the total exciton wave-function has Γ_8^- symmetry.

The actual determination of the d-type matrices can be carried out by standard group theoretical methods involving the use of projection operators, or by a simpler method described by SCHECHTER[7]. For the radial parts of the trial functions we take $\exp(-r/r_1)$ for the s-terms and $\exp(-r/r_2)$ for the d-terms. We consider the algebraic form of the trial functions separately for the three cases in which we are interested.

(a) *The direct exciton in germanium*

The number of d-terms in the trial function has been determined above, and the complete function is found to be:

$$c_1\begin{bmatrix}1\\0\\0\\0\end{bmatrix}\exp(-r/r_1)+c_2\begin{bmatrix}z^2-\frac{1}{2}(x^2+y^2)\\0\\\sqrt{(3/2)}(x^2-y^2)\\0\end{bmatrix}\exp(-r/r_2)+$$

$$+c_3\begin{bmatrix}0\\iz(x+iy)\\ixy\\0\end{bmatrix}\exp(-r/r_2)+c_4\begin{bmatrix}0\\iz(x+iy)\\-2ixy\\i\sqrt{(3)}z(x-iy)\end{bmatrix}\exp(-r/r_2).$$

Here the four c's are variational parameters and also r_1 and r_2. The last matrix in the sum is not coupled to the s-matrix by the exciton Hamiltonian, but only to the other d-matrices, so that its effect on the exciton binding energy is small. There are, of course, three other functions similar to the above, but since they are degenerate with it they need not be considered.

We note that an exciton state of this type can be produced optically, for the matrix element governing such transitions contains the ground-state wave-function of symmetry Γ_1 and the gradient operator transforming as Γ_{15}, so that the final state wave-function must transform as $\Gamma_1 \times \Gamma_{15} = \Gamma_{15}$ for the matrix element to be non-zero. The electron spin function must now be included and transforms like $D_{1/2}$ of the rotation group, so that the total exciton wave-function has symmetry $\Gamma_8^- \times D_{1/2} = \Gamma_{12}' + \Gamma_{15} + \Gamma_{20}$. This contains Γ_{15}, so that optical transitions into this exciton state are allowed.

(b) *The indirect exciton in germanium*

We have shown above that the products of valence- and conduction-band Bloch functions in this case form a set of basis functions for $L_4^+(\boldsymbol{k}_c) + L_5^+(\boldsymbol{k}_c) + L_6^+(\boldsymbol{k}_c)$. The angular-momentum d-functions can be used to form a basis for the irreducible representations $L_1(0) + 2L_3(0)$. The products of these two sets of functions therefore form bases for:

$$[L_1(0)+2L_3(0)] \times [L_4^+(\boldsymbol{k}_c)+L_5^+(\boldsymbol{k}_c)+L_6^+(\boldsymbol{k}_c)] = 3[L_4^+(\boldsymbol{k}_c)+L_5^+(\boldsymbol{k}_c)]+7L_6^+(\boldsymbol{k}_c).$$

There are therefore three matrices of d-symmetry which can be included in the trial function for the $L_4^+(\boldsymbol{k}_c)+L_5^+(\boldsymbol{k}_c)$ exciton and seven such matrices for the case of the $L_6^+(\boldsymbol{k}_c)$ exciton. We shall not exhibit here the actual form of the trial functions, but will confine ourselves to giving the result of the variational calculation below. It is easily verified that both these exciton states can be excited by indirect transitions involving either of the two acoustical phonons at the L symmetry point, but not the optical phonons.

(c) *The indirect exciton in silicon*

(i) *Infinite spin-orbit splitting.* There are two twofold degenerate exciton levels having symmetries $\Delta_6(\boldsymbol{k}_c)$ and $\Delta_7(\boldsymbol{k}_c)$. The angular momentum d-functions can be used to form a basis for $\Delta_1(0)+\Delta_2(0)+\Delta_2'(0)+\Delta_5(0)$. We have:

$$[\Delta_1(0)+\Delta_2(0)+\Delta_2'(0)+\Delta_5(0)] \times [\Delta_6(\boldsymbol{k}_c)+\Delta_7(\boldsymbol{k}_c)] = 5\Delta_6(\boldsymbol{k}_c)+5\Delta_7(\boldsymbol{k}_c)$$

so that for each exciton there are five d-matrices having the correct symmetry properties for inclusion in the trial function. Either of the two types of exciton level can be excited by indirect transitions involving any type of phonon in the Δ symmetry direction.

(ii) *Zero spin-orbit splitting.* The two exciton levels are a singlet $\Delta_2'(\boldsymbol{k}_c)$ and a doublet $\Delta_5(\boldsymbol{k}_c)$. Now:

$$[\Delta_1(0)+\Delta_2(0)+\Delta_2'(0)+\Delta_5(0)] \times [\Delta_2'(\boldsymbol{k}_c)+\Delta_5(\boldsymbol{k}_c)] = 2\Delta_1(\boldsymbol{k}_c)+2\Delta_1'(\boldsymbol{k}_c)+\Delta_2(\boldsymbol{k}_c)+2\Delta_2'(\boldsymbol{k}_c)+4\Delta_5(\boldsymbol{k}_c)$$

so there are two d-type matrices for the $\Delta_2'(\boldsymbol{k}_c)$ exciton trial function and four such matrices for the $\Delta_5(\boldsymbol{k}_c)$ exciton. Either exciton level can be excited by indirect transitions involving any type of phonon in the Δ symmetry direction, with the one exception that the $\Delta_2'(\boldsymbol{k}_c)$ level cannot be excited by a transition involving the longitudinal acoustic phonon.

4. RESULTS

Minimizing the energies, corresponding to the various trial functions, with respect to the parameters c_i, $i = 1,2 \ldots n$, r_1 and r_2 involves finding the minimum latent root of a certain $n \times n$ determinant (whose elements are functions of r_1 and r_2) which can be produced by variation of r_1 and r_2. This minimization procedure was carried out numerically in each case on the R.R.E. electronic computer. The results are shown in Table 1 for germanium and Table 2 for silicon.

For the direct exciton in germanium the calculated binding energy of 0·0014 eV is only about 10 per cent larger than the value obtained by taking only the s-part of the trial function, so that one would expect the inclusion of higher angular-momentum states than d-states to produce a negligible increase in the binding energy. This value agrees reasonably with the R.R.E. experimental value of $0{\cdot}0011 \pm 0{\cdot}0001$ eV, but is in disagreement with the Lincoln Laboratory value of $0{\cdot}0025 \pm 0{\cdot}0004$ eV. However, both of these values may be in error; the R.R.E.

Table 1. *The binding energies of excitons in germanium, using both s-type and (s+d)-type trial functions and band parameters quoted by* KOHN[5], *and the corresponding values of the variational parameters* r_1 *and* r_2.

	Binding energy (eV)	r_1 (Å)	r_2 (Å)
Direct exciton—*s*-type function	0·00127	350	—
Direct exciton—*s*- and *d*-type functions	0·00138	320	220
Indirect exciton—*s*-type function	0·00244	180	—
Indirect exciton—*s*- and *d*-type functions ($L_4^+ + L_5^+$)	0·00347	120	87
Indirect exciton—*s*- and *d*-type functions L_6^+	0·00288	150	85

VOL. 13 1960

value was obtained from an analysis of the absorption spectrum appropriate to a material having simple band edges, the error quoted here merely being due to a scatter in the values obtained at different temperatures; the Lincoln Laboratory value was obtained more directly, but from a specimen in a state of strain which has been shewn to influence considerably the value of the energy gap.[16] This strain resulted from the difference in thermal contraction of the thin slab of germanium and the glass mounting to which it was cemented. On cooling to liquid-helium temperatures, a uniform isotropic strain in the plane of the germanium slab is thus set up. In the presence of a strain of this type, the valence-band degeneracy at $\boldsymbol{k} = 0$ is lifted, giving two valence-band edges and correspondingly two groups of exciton states. At sufficiently small values of $\boldsymbol{k}$, the surfaces of constant energy for the two valence bands are ellipsoids of revolution about the direction of the slab normal, being prolate ellipsoids for one band and oblate ellipsoids for the other. If we assume that

Table 2. *The binding energies of the indirect excitons in silicon, using both s and (s+d)-type trial functions and band parameters quoted by* KOHN[5], *and the corresponding values of the variational parameters* r_1 *and* r_2 *for both zero and infinitely large spin-orbit coupling and for both positive and negative values of the effective mass parameter* $(L-M)$.[15]

				Binding energy (eV)	r_1 (Å)	r_2 (Å)
Zero s.o. coupling	*s*-type function	$(L-M) \lesssim 0$		0·0120	50	—
	s- and *d*-type functions	$(L-M)<0$	Δ_2'	0·0132	44	33
			Δ_5	0·0137	42	30
	s- and *d*-type functions	$(L-M)>0$	Δ_2'	0·0145	39	28
			Δ_5	0·0133	44	30
Infinite s.o. coupling	*s*-type function	$(L-M) \lesssim 0$		0·0120	50	—
	s- and *d*-type functions	$(L-M)<0$	Δ_6	0·0130	44	31
			Δ_7	0·0124	47	33
	s- and *d*-type functions	$(L-M)>0$	Δ_6	0·0124	47	33
			Δ_7	0·0130	44	31

the splitting between the bands at $\boldsymbol{k} = 0$ is large compared with the exciton binding energy, then it is this region of ellipsoidal energy-dependence which is important in determining the hole effective mass for use in the exciton wave-equation. Taking the conduction-band effective mass at $\boldsymbol{k} = 0$ to be unaltered by the strain, the exciton functions $F_{no}(\boldsymbol{r})$ then satisfy an equation similar to that for a donor state in germanium or silicon, and the ground-state binding energies can be calculated by the same method as has been used for the donor problem.[17] It is found that the excitons due to both bands have approximately the same binding energy of about 0·0013 eV. Since the measured value of the valence-band splitting in the Lincoln Laboratory experiments was 0·005 eV,[18] the exciton binding energy is as large as one-quarter of the splitting, so that the assumption on which the above estimate is based is not very well satisfied. However, the calculation indicates that the difference between the binding energies observed by the two experimental groups is not likely to be entirely accounted for by the strain in the Lincoln Laboratory specimen.

As for the accuracy of the calculated binding energy, we may note that a similar theory applied to the ground states of acceptors in germanium[7] gives a binding energy which is about 20 per cent smaller than that measured experimentally, and the discrepancy is thought to be largely due to breakdown of effective-mass theory in the immediate vicinity of the impurity atom. However, the mean radius of the direct exciton wave-function in germanium is almost ten times that of the ground-state acceptor wave-function, so that effects due to the breakdown of effective-mass theory when the electron and hole are close together would be expected to be correspondingly smaller. Further, for the exciton there is no distortion of the lattice analogous to that introduced by the inclusion of the impurity atom which renders the effective-mass theory less accurate for the case of an acceptor. A better agreement between the calculated and experimental results for the direct exciton in germanium might therefore be expected.

For the indirect exciton in germanium, the calculated ground-state binding energy is 0·0035 eV, and an excited state lies above this by an energy 0·0006 eV. The two binding energies represent substantial increases over the values obtained by using only an *s*-type trial function. Experimentally, the R.R.E. value for the ground-state binding energy is $0{\cdot}0027 \pm 0{\cdot}0004$ eV, with an excited state at $0{\cdot}0010 \pm 0{\cdot}0001$ eV above this, while for the same quantities Lincoln Laboratory measures $0{\cdot}0033 \pm 0{\cdot}0004$ and 0·0010 eV, respectively. The reported ground-state binding energies are therefore consistent with a value of about 0·0030 eV, and the splittings between the two levels agree at 0·0010 eV. There is fair agreement between experimental and theoretical values for the ground-state binding energy, but for the splitting between the two levels the calculated value is significantly lower than that measured experimentally. We note that the energies calculated with the present trial functions are larger than the values of 0·0025 and 0·0033 eV calculated by the Lincoln Laboratory,[2, 10] using the trial function $\exp\{-[a^2z^2+b^2(x^2+y^2)]\}$.

For the various cases considered for the indirect exciton in silicon, the inclusion of the *d*-states in the trial function produces at most an increase of about 20 per cent over the binding energy found by using only an *s*-type trial function. The calculated ground-state binding energy is about 0·012–0·014 eV, and the excited state lies above this by about 0·0005–0·0012 eV. These results are to be compared with the values measured at R.R.E. of about 0·01 eV for the binding energy, and 0·0055 eV for the splitting. The agreement between the calculated and observed ground-state binding energies is fair considering the uncertainty in the latter, but the agreement between the splittings is so bad as to suggest that the excited state observed experimentally is not that whose energy we have calculated.

The mean radii of the indirect exciton wave-functions are smaller, particularly in silicon, than that of the direct exciton in germanium, and, on the basis of the above remarks concerning the breakdown of effective-mass theory when the electron and hole are close together, we should expect the calculated values for these indirect exciton binding energies to be less accurate than the calculated binding energy of the direct exciton in germanium. However, the mean indirect exciton radii are larger than the ground acceptor state radii in the same semiconductor by a factor of about three, so that the accuracy of the exciton calculations would be expected to be correspondingly better.

For each of the three types of exciton there is fair agreement between the measured and calculated values of the ground-state binding energies. A more detailed comparison of theory with experiment must await more accurate experiments on silicon and a reconciliation of the two experimental groups whose differing observations of the direct exciton in germanium we have discussed.

Acknowledgements—The authors wish to thank the Lincoln Laboratory for pre-publication copies of their results. Their thanks are also due to Dr. D. P. JENKINS for help in programming some parts of the calculation for the computer.

This work was principally carried out in the summer of 1958, when one of us (R.L.) was a Vacation Student at the Royal Radar Establishment. He would like to take this opportunity to thank Dr. G. G. MACFARLANE and his group for a pleasant and stimulating time.

This paper is published by permission of the Controller, Her Britannic Majesty's Stationery Office.

VOL. 13 1960

REFERENCES

1. MACFARLANE G. G., MCLEAN T. P., QUARRINGTON J. E. and ROBERTS V., *Phys. Rev.* **108**, 1377 (1957); *Proc. Phys. Soc. Lond.* **71**, 863 (1958); *Phys. Rev.* **111**, 1245 (1958); *J. Phys. Chem. Solids* **8**, 388 (1959).
2. ZWERDLING S., ROTH L. M. and LAX B., *Phys. Rev.* **109**, 2207 (1958); *J. Phys. Chem. Solids* **8**, 397 (1959).
   BUTTON K. J., ROTH L. M., KLEINER W. H., ZWERDLING S. and LAX B., *Phys. Rev. Letters* **2**, 161 (1959).
3. DRESSELHAUS G., *J. Phys. Chem. Solids* **1**, 14 (1956).
4. WANNIER G. H., *Phys. Rev.* **52**, 191 (1937).
5. KOHN W., *Solid State Physics* **5**, 257 (1957).
6. KOHN W. and SCHECHTER D., *Phys. Rev.* **99**, 1903 (1955).
7. SCHECHTER D., Thesis, Carnegie Institute of Technology (1958).
8. KOSTER G. F., *Solid State Physics*, 5, 173 (1957).
9. KAHN A. H., *Phys. Rev.* **97**, 1647 (1955).
10. ROTH L. M., LAX B. and ZWERDLING S., *Phys. Rev* to be published.
11. ELLIOTT R. J., *Phys. Rev.* **96**, 266 (1954).
12. HERMAN F., *Phys. Rev.* **95**, 847 (1954).
13. BOUCKAERT L. P., SMOLUCHOWSKI R. and WIGNER E., *Phys. Rev.* **50**, 58 (1936).
14. ELLIOTT R. J., *Phys. Rev.* **96**, 280 (1954).
15. DRESSELHAUS G., KIP A. F. and KITTEL C., *Phys. Rev.* **98**, 368 (1955).
16. MACFARLANE G. G., MCLEAN T. P., QUARRINGTON J. E. and ROBERTS V., *Phys. Rev. Letters* **2**, 252 (1959).
17. KOHN W. and LUTTINGER J. M., *Phys. Rev.* **98**, 915 (1955).
18. KLEINER W. H. and ROTH L. M., *Phys. Rev. Letters* **2**, 334 (1959).

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 13. pp. 10–27. Printed in Great Britain.

FLUCTUATIONS d'AIMANTATION ET OPALESCENCE CRITIQUE

P. G. de GENNES et J. VILLAIN

CEN de Saclay, B.P. 2 — Gif s/Yvette (S. & O.), France

(*Received* 22 *April* 1959)

Résumé—On étudie les fluctuations d'aimantation des substances où plusieurs types d'ions magnétiques sont simultanément présents.

On détermine d'abord au moyen d'une forme locale de l'approximation de champ moléculaire, l'intensité des fluctuations correspondant à chaque vecteur d'onde. On montre en particulier que les résultats obtenus au-dessous du point de transition se raccordent exactement, aux basses températures, avec les formules rigoureuses déduites de la théorie des ondes de spin. On envisage ensuite la relaxation des fluctuations dans le domaine paramagnétique, et plus particulièrement les phénomènes de "relaxation lente".

On retrouve les phénomènes connus d'opalescence critique, mais avec des anomalies de comportement qui ne se rencontrent pas dans les cas les plus simples envisagés antérieurement. Certaines de ces anomalies sont observables par diffusion magnetique des neutrons.

Le calcul de ces effets est présenté en détail pour la magnétite, où l'on doit décomposer les ions magnétiques en six sous réseaux. On montre également que pour l'opalescence critique des antiferromagnétiques, la relaxation est lente, et la distribution de fréquences Lorentzienne: ceci provient d'un effet de freinage thermodynamique, et ne se relie pas à une diffusion de spin comme dans les ferromagnétiques.

Abstract—This paper deals with fluctuations in magnetization in substances in which several types of magnetic ions are present simultaneously.

First, by means of a localized form of the molecular field approximation, the intensity of fluctuations corresponding to each wave vector is studied. It is shown, in particular, that the results obtained below the transition temperature agree exactly at low temperatures with the rigorous formulae deduced from the theory of spin waves. Next, the relaxation of fluctuations in the paramagnetic range is considered, particularly the phenomena of "slow relaxation".

Critical opalescence phenomena are found, but they show important differences in behavior from what is known from the simple situations considered up to now. Some of these anomalies may be observed in neutron-scattering experiments. These effects are worked out numerically for magnetite, where six non-equivalent sublattices have to be considered. It is also shown that for antiferromagnets the relaxation of the critical fluctuations is slow, and, that their frequency spectrum is of the Lorentz type. This is a thermodynamic effect and is distinct from the spin diffusion effect familiar in ferromagnets.

1. INTRODUCTION

DES anomalies de la diffraction des neutrons au voisinage de certaines réflexions de Bragg ont été mises en évidence dans différentes substances ferro-[1–6], antiferro-[7–8] ou ferrimagnétiques[9–10] pour des températures proches du point de transition.

L'analogie entre ces phénomènes et l'opalescence critique des fluides a été remarquée par VAN HOVE[11], et il les a attribués à des fluctuations d'aimantation. Les paramètres phénoménologiques introduits par lui à cette occasion (pour un cristal ferromagnétique cubique à 1 spin par maille) ont été calculés ensuite[12–15] moyennant certaines approximations de nature statistique; toutefois, lorsque l'on en vient à des situations plus complexes, la disparition de certaines symétries fait que l'analyse phénoménologique de VAN HOVE doit être modifiée de façon non triviale. Nous avons noté ce fait pour la première fois à

VOL. 13 1960

propos de certains antiferromagnétiques cubiques tels que MnO.[15] A la suite des expériences de RISTE *et al.* sur la magnétite,[10] il nous a paru utile de discuter plus généralement l'opalescence critique magnétique dans les réseaux complexes, puis d'illustrer ces considérations par l'exemple de Fe_3O_3. Pour pousser les calculs jusqu'aux applications pratiques, nous aurons fréquemment recours à :

(*a*) une approximation semi-classique,

(*b*) une version locale de l'approximation de champ moléculaire (seule praticable pour les cas qui nous intéressent ici).

$$\mathscr{H} = -\sum_{\substack{ij \\ R_i R_j}} J_{R_i R_j}\, \boldsymbol{S}_{R_i} \cdot \boldsymbol{S}_{R_j} \qquad (1$$

Nous négligeons complètement l'effet de la relaxation spin réseau sur la cinétique des fluctuations.

Le modèle physique étant ainsi précisé, nous rappellerons brièvement comment la section efficace de diffusion des neutrons se relie aux fluctuations.

Si l'on néglige l'effet du magnétisme orbital, cette section efficace différentielle par élément d'angle solide $d\Omega$ s'écrit[11]

$$\frac{d\sigma}{d\Omega d\omega} = \left(\frac{1{,}91e^2}{mc^2}\right)^2 \frac{k_f}{k_0} \sum_{ij} F_i(\boldsymbol{k}) F_j(\boldsymbol{k}) \int \frac{dt}{2\pi} \exp(-i\omega t) \sum_{R_i R_j} \sum_{\epsilon\eta} \langle S^{\epsilon}_{R_i}(0) S^{\eta}_{R_j}(t) \rangle \exp\left[i\boldsymbol{k}(\boldsymbol{R}_i - \boldsymbol{R}_j)\right] \left(\delta_{\epsilon\eta} - \frac{k_\epsilon k_\eta}{k^2}\right) \qquad (2)$$

Du point de vue physique le modèle adopté est celui de Heisenberg : les spins $\boldsymbol{S}_{R_i}$ sont localisés aux noeuds $\boldsymbol{R}_i$ d'un réseau.

La maille de référence contient n ions magnétiques, de coordonnées

$$\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_n.$$

La maille que nous adoptons ici est la maille "magnétique", et elle peut être plus grande parfois que la maille cristalline. Cette convention permet d'unifier l'exposé relatif aux températures supérieures ou inférieures au point de transition. (Pour un calcul explicite sur un exemple, il peut être plus simple de s'en tenir à la maille cristalline comme dans la réf 15.

Les coordonnées les plus générales des ions magnétiques sont :

$$\boldsymbol{R}_i = \boldsymbol{r}_i + \nu_a \boldsymbol{a} + \nu_b \boldsymbol{b} + \nu_c \boldsymbol{c}$$

où $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ définissent le groupe de translation du cristal ; $\nu_a\ \nu_b\ \nu_c$ sont des entiers.

Nous supposons que les spins sont couplés entre aux uniquement par une énergie d'échange

Dans cette formule $\hbar\boldsymbol{k}_0$ et $\hbar\boldsymbol{k}_f$ sont les impulsions initiales et finales du neutron. L'énergie et l'impulsion transférées lors du choc sont respectivement

$$\hbar\omega = \frac{\hbar^2}{2M}(k_0^2 - k_f^2) \text{ et } \hbar\boldsymbol{k} = \hbar(\boldsymbol{k}_0 - \boldsymbol{k}_f).$$

M est la masse du neutron, e^2/mc^2 le rayon classique de l'électron, $F_i(\boldsymbol{k})$ le facteur de forme magnétique de l'ion (i), normalisé pour que $F_i(0) = 1$. Le symbole $\langle \ldots \rangle$ représente une moyenne thermique sur les niveaux du système de spin.

Nous poserons

$$T_{j,\boldsymbol{k}} = \sum_{R_j} \boldsymbol{S}_{R_j} \exp[i\boldsymbol{k}(\boldsymbol{R}_j - \boldsymbol{r}_j)] \qquad (3)$$

(le facteur de phase étant choisi tel que $\boldsymbol{T}_{j,\boldsymbol{k}}$ ait en fonction de $\boldsymbol{k}$ la périodicité du réseau réciproque).

La section efficace peut alors être récrite

$$\frac{d\sigma}{d\Omega d\omega} = \left(\frac{1{,}91e^2}{mc^2}\right)^2 \frac{k_f}{k_0} \sum_{ij} F_i(\boldsymbol{k}) F_j(\boldsymbol{k}) \int \frac{dt}{2\pi} \exp(-i\omega t) \sum_{\epsilon\eta} \langle T^{\epsilon}_{i\boldsymbol{k}}(0) T^{\eta}_{j,-\boldsymbol{k}}(t) \rangle \exp\left[i\boldsymbol{k}(\boldsymbol{r}_i - \boldsymbol{r}_j)\right] \left(\delta_{\epsilon\eta} - \frac{k_\epsilon k_\eta}{k^2}\right) \qquad (4)$$

On peut séparer dans cette formule :

(*a*) un terme de diffusion strictement élastique dû à l'ordre à longue distance des spins, non nul

seulement en-dessous du point de transition T_c et dans les directions de Bragg,

(*b*) un terme (généralement inélastique) dû aux fluctuations d'aimantation, qui nous intéresse seul ici, et qui s'écrit explicitement :

$$\left(\frac{d\sigma}{d\Omega d\omega}\right)_{\text{fluct}} = \left(\frac{1{,}91\ e^2}{mc^2}\right)^2 \frac{k_f}{k_0} \sum_{ij} F_i(\boldsymbol{k}) F_j(\boldsymbol{k}) \int \frac{dt}{2\pi} \sum_{\epsilon\eta} \left[\langle T^{\epsilon}_{i\boldsymbol{k}}(0) T^{\eta}_{j,-\boldsymbol{k}}(t) \rangle - \langle T^{\epsilon}_{i-\boldsymbol{k}} \rangle \langle T^{\eta}_{j,\boldsymbol{k}} \rangle \right]$$

$$\exp i\boldsymbol{k}(\boldsymbol{r}_i - \boldsymbol{r}_j) \left(\delta_{\epsilon\eta} - \frac{k_\epsilon k_\eta}{k^2}\right) \exp(-i\omega t) \qquad (5)$$

C'est ce terme que nous allons étudier plus en détail dans les sections suivantes.

2. APPROXIMATION QUASI-STATIQUE

Lorsque l'énergie du neutron incident est grande devant la moyenne des transferts $\hbar\omega$, on peut écrire directement la section efficace intégrée sur les énergies finales du neutron sous la forme :

$$\left(\frac{d\sigma}{d\Omega d\omega}\right)_{\text{fluct}} = \left(\frac{1{,}91\ e^2}{mc^2}\right)^2 \sum_{\substack{i,j \\ \epsilon,\eta}} F_i(\boldsymbol{k}) F_j(\boldsymbol{k}) \langle (T^{\epsilon}_{i\boldsymbol{k}} - \langle T^{\epsilon}_{i\boldsymbol{k}} \rangle)(T^{\eta}_{j,-\boldsymbol{k}} - \langle T^{\eta}_{j,-\boldsymbol{k}} \rangle) \rangle \exp i\boldsymbol{k}(\boldsymbol{r}_i - \boldsymbol{r}_j) \left(\delta_{\epsilon\eta} - \frac{k_\epsilon k_\eta}{k^2}\right) \qquad (6)$$

Dans (6) toutes les variables T sont prises au même instant, ce qui simplifie radicalement le calcul des moyennes. Nous allons maintenant relier ces moyennes à un système de susceptibilités pour une perturbation magnétique extérieure de vecteur d'onde $\boldsymbol{k}$.

Nous suivons la méthode de perturbation d'Yvon[16] déjà appliquée précédemment à des réseaux simples de spins.[15]

A chaque site $\boldsymbol{R}_j$ nous appliquons un champ extérieur $\boldsymbol{H}_{R_j}$; l'ion de coordonnée $\boldsymbol{R}_j$ porte alors un moment magnétique moyen

$$\mu_{R_i} = g\mu_B \frac{\operatorname{Tr}\{\boldsymbol{S}_{R_i} \exp[-\beta(\mathscr{H}_0 + \delta\mathscr{H})]\}}{\operatorname{Tr} \exp -\beta[(\mathscr{H}_0 + \delta\mathscr{H})]}$$

$$\delta\mathscr{H} = -g\mu_B \sum_{j,R_j} \boldsymbol{S}_{R_j} . \boldsymbol{H}_{K_j} \qquad (7)$$

g est le facteur de séparation spectroscopique égal à 2 puisque nous supposons le moment orbital bloqué, et μ_B le magnéton de Bohr électronique. $\beta = 1/k_B T$.

Dans toute la suite nous envisagerons seulement la *limite classique* des propriétés de notre assemblée de spins. On peut s'attendre, à priori, à ce que cette approximation soit acceptable lorsque la longueur S des spins est notablement supérieure à 1 (en continuant toutefois à écrire

$$S_i^2 = S_i(S_i + 1)$$

pour améliorer la précision). Un critère plus détaillé de validité est obtenu en écrivant que pour toutes les fréquences ω qui sont contenues dans le spectre de puissance des fluctuations considerées on a :

$$\hbar\omega < k_B T$$

Dans le domaine d'opalescence critique la relaxation des fluctuations intéressantes est lente (pour différentes raisons que nous étudierons plus loin). Le spectre de puissance est alors étroit, et la condition ci-dessus est satisfaite (même pour $S = \frac{1}{2}$). Cette remarque étant faite, le calcul de μ_{R_j} peut être poursuivi en négligeant dans (7) la non commutation de l'hamiltonien d'échange $\mathscr{H}_0$ avec $\delta\mathscr{H}$. Si nous appelons $\delta\mu_{R_j}$ la différence entre μ_{R_j} et sa valeur d'équilibre en champ nul, nous obtenons au premier ordre en H :

$$\delta\mu^{\epsilon}_{R_i} = \beta g^2 \mu_B^2 \sum_{j,R_j} \sum_{\eta} \left[\frac{\operatorname{Tr}\{\exp(-\beta\mathscr{H}_0) S^{\epsilon}_{R_i} S^{\eta}_{R_j}\}}{Z} H^{\eta}_{R_j} - \right.$$

$$\left. - \frac{\operatorname{Tr}\{\exp(-\beta\mathscr{H}_0) S^{\epsilon}_{R_i}\} \operatorname{Tr}\{\exp(-\beta\mathscr{H}_0) S^{\eta}_{R_j}\}}{Z^2} H^{\eta}_{R_j} \right] \qquad (8)$$

VOL. 13 1960

où $Z = \mathrm{Tr}\exp(-\beta\mathscr{H})$.

$$\delta\mu^{\epsilon}_{R_i} = \beta g^2\mu_B^2 \sum_{j,R_j}\sum_{\eta}\Big[\langle S^{\epsilon}_{R_j}S^{\eta}_{R_i}\rangle - \langle S^{\epsilon}_{R_j}S^{\eta}_{R_i}\rangle\Big]H^{\eta}_{R_j} \tag{9}$$

Introduisons maintenant les transformées de Fourier

$$\mu_{ik} = g\mu_B T_{ik}$$

$$H_{i,k} = \sum H_{R_i}\exp\left[i\mathbf{K}(\mathbf{R}_i - r_i)\right] \tag{10}$$

et définissons un système de susceptibilités par les relations

$$N\mu^{\epsilon}_{ik} = \sum_j\sum_{\eta}\chi^{\epsilon\eta}_{ij}(\mathbf{K})H^{\eta}_{jk} \tag{11}$$

(où N est le nombre de mailles du spécimen).

L'équation (9) relie ces susceptibilités aux corrélations,

$$\chi^{\epsilon\eta}_{ij}(\boldsymbol{k}) = \beta g^2\mu_B^2\sum_{R_iR_j}[\langle S^{\epsilon}_{R_i}S^{\eta}_{R_j}\rangle - \langle S^{\epsilon}_{R_i}\rangle\langle S^{\eta}_{R_j}\rangle]\exp\left[i\boldsymbol{k}(\boldsymbol{R}_i - \boldsymbol{r}_i) - i\boldsymbol{k}(\boldsymbol{R}_j - \boldsymbol{r}_j)\right] \tag{12}$$

$$\chi^{\epsilon\eta}_{ij}(\boldsymbol{k}) = \beta g^2\mu_B^2[\langle T^{\epsilon}_{ik}T^{\eta}_{j,-k}\rangle - \langle T^{\epsilon}_{ik}\rangle\langle T^{\eta}_{j,-k}\rangle]. \tag{13}$$

Cette dernière formule permet de relier la section efficace dans l'approximation quasi-statique (6) aux susceptibilités que nous venons de définir

$$\left(\frac{d\sigma}{d\Omega}\right)_{\mathrm{fluct}} = \left(\frac{1{,}91e^2}{mc^2}\right)\frac{k_BT}{g^2\mu_B^2}\times \sum_{ij}\sum_{\epsilon\eta}F_i(\boldsymbol{k})F_j(\boldsymbol{k})\chi^{\epsilon\eta}_{ij}(\boldsymbol{k})\left[\delta_{\epsilon\eta} - \frac{k_\epsilon k_\eta}{k^2}\right]\exp[i\boldsymbol{k}(\boldsymbol{r}_i - \boldsymbol{r}_j)] \tag{14}$$

Enfin les susceptibilités $\chi^{\epsilon\eta}_{ij}(\boldsymbol{k})$ peuvent être calculées à partir des intégrales d'échange en écrivant que chaque spin est en équilibre sous l'action du champ extérieur et du champ moléculaire du à ses voisins. Nous conduirons ce calcul successivement pour les températures supérieures et inférieures au point de transition.

FLUCTUATIONS DANS LE DOMAINE PARAMAGNÉTIQUE

Au-dessus du point de transition, une forme locale de l'équation de champ moléculaire s'écrit :

$$\boldsymbol{\mu}_{R_i} = \frac{S_i(S_i+1)}{3k_BT}\left[g^2\mu_B^2\boldsymbol{H}_{R_i} + \sum_{jR_j}2J_{R_iR_j}\boldsymbol{\mu}_{R_j}\right] \tag{15}$$

Nous définirons une matrice A hermitique, de rang n, pour chaque vecteur d'onde $\boldsymbol{k}$ par la relation

$$A_{ij}(\boldsymbol{k}) = \delta_{ij}\frac{3k_BT}{S_i(S_i+1)} - \sum_{R_j}2J_{R_iR_j}\times \exp\left[i\boldsymbol{k}\cdot(\boldsymbol{R}_i - \boldsymbol{R}_j - \boldsymbol{r}_i + \boldsymbol{r}_j)\right] \tag{16}$$

La relation (15) peut être récrite

$$\sum A_{ij}(\boldsymbol{k})\mu_{jk} = g^2\mu_B^2H_{ik} \tag{17}$$

ou, en faisant intervenir la matrice inverse A^{-1}, et en comparant avec (11)

$$\chi^{\epsilon\eta}_{ij}(\boldsymbol{k}) = Ng^2\mu_B^2[A^{-1}(\boldsymbol{k})]_{ij}\delta_{\epsilon\eta} \tag{18}$$

$$\left(\frac{d\sigma}{d\Omega}\right)_{\mathrm{fluct}} = 2N\left(\frac{1{,}91e^2}{mc^2}\right)^2k_BT\times \sum F_i(\boldsymbol{k})F_j(\boldsymbol{k})[A^{-1}(\boldsymbol{k})]_{ij}\exp\left[i\boldsymbol{k}(\boldsymbol{r}_i - \boldsymbol{r}_j)\right] \tag{19}$$

Pour chaque vecteur d'onde le calcul de la section efficace est ainsi ramené à l'inversion d'une matrice A de rang n. Nous poserons explicitement

$$[A^{-1}(\boldsymbol{k})]_{ij} = \frac{\tilde{A}_{ij}(\boldsymbol{k})}{\Delta(\boldsymbol{k})} \tag{20}$$

où Δ est le déterminant de la matrice A, et $\tilde{A}_{ji}$ le mineur relatif à l'élément A_{ij}. Toutes ces quantités sont des fonctions de $\boldsymbol{k}$ qui ont la périodicité du réseau réciproque. Δ est un polynome de degré n en T. $\tilde{A}_{ij}$ est un polynome de degré $n-1$ en T.

Chaque terme (ij) de la section efficace (19) comprend

(a) deux facteurs de forme,

(b) une fonction $[A^{-1}(\boldsymbol{k})_i]_j$ qui a la périodicité du réseau réciproque,

(c) le facteur $\exp[i\boldsymbol{k}(\boldsymbol{r}_i - \boldsymbol{r}_j)]$ qui n'est pas en général périodique dans le réseau réciproque.

A chaque température T il est particulièrement intéressant de déterminer les vecteurs $\boldsymbol{k}$ pour lesquels la diffusion est maximum. Ce problème dépend simultanément des facteurs (*a*), (*b*), (*c*). Nous discuterons plus particulièrement le cas où es facteurs de forme des différents ions ont même dépendance en $\boldsymbol{k}$.

Leur rôle est alors trivial, et la section efficace est, pour chaque $\boldsymbol{k}$, proportionnelle à

$$I(\boldsymbol{k}) = \frac{3k_BT\sum_{ij}\bar{A}_{ij}(\boldsymbol{k})\exp[i\boldsymbol{k}(\boldsymbol{r}_i-\boldsymbol{r}_j)]}{\Delta(\boldsymbol{k})\sum_i S_i(S_i+1)} \qquad (20)$$

La normalisation de I est telle que, à température, infinie, $I(\boldsymbol{k}) \equiv 1$, les propriétés générales de $I(\boldsymbol{k})$ sont les suivantes :

(1) $I(\boldsymbol{k})$ est extremum pour $\boldsymbol{k} = 0$, mais pas, en général, sut les autres noeuds du réseau réciproque. Si toutefois les coordonnées $\boldsymbol{r}_i$ des atomes magnétiques ne dépendent d'aucun paramètre, les facteurs $\exp[i\boldsymbol{k}(\boldsymbol{r}_i - \boldsymbol{r}_j)]$ ont une période commune multiple de celle du réseau réciproque. $I(\boldsymbol{k})$ est alors extremum sur les noeuds du réseau réciproque $\boldsymbol{\tau}$ séparés de l'origine par un nombre entier de ces périodes. Nous dirons que de tels noeuds sont équivalents à $\boldsymbol{k} = 0$. Ils correspondent à un maximum de $I(\boldsymbol{k})$ pour les substances qui, en-dessous du point de transition, sont ferromagnétiques et à un minimum pour les antiferromagnétiques (pour les antiferromagnétiques du type MnF_2 ou MnO, $I(\boldsymbol{k})$ coïncide avec la quantité $\chi(\boldsymbol{k})/\chi_1$ définie et étudiée dans la réf (15). Nous verrons plus loin que pour une ferrite comme la magnétite, les noeuds équivalents à O correspondent à un maximum relatif de $I(\boldsymbol{k})$.

D'une façon générale, si $I(\boldsymbol{k})$ est maximum sur les noeuds équivalents à O, et si le cristal est cubique on peut caractériser la diffusion au voisinage de ces points par deux paramètres $r_1(o)$ et κ_1 définis par

$$\frac{1}{I(\boldsymbol{k})} \cong r_1^2(0)[\kappa_1^2+(\boldsymbol{k}-\boldsymbol{\tau})^2] \qquad (21)$$

$$[k_1r_1(0)]^2\frac{1}{I(0)} = \frac{\sum_i \chi_0^i}{\sum_{ij}\chi_{ij}(0)} \qquad (22)$$

en complète analogie avec la notation de VAN HOVE[11]. Les χ_0^i désignent les susceptibilités des sous réseaux en l'absence d'interactions d'échange. On vérifie facilement que, pour toutes les substances autres que les antiferromagnétiques, κ_1^2 s'annule à T_c et croit linéairement avec la température au voisinage supérieur de T_c.

(2) Etudions maintenant les conditions générales de l'opalescence critique, c'est à dire de l'apparition de valeurs très grandes de $I(\boldsymbol{k})$ pour T voisin de T_c.

Une valeur infinie de I ne peut être obtenue que si le déterminant Δ s'annule. Ce déterminant est positif dans tout le domaine paramagnétique, quel que soit $\boldsymbol{k}$. Il s'annule pour $T = T_c$ si $\boldsymbol{k}$ coïncide avec certains *vecteurs critiques* $\boldsymbol{k}_c$: toutes les directions d'opalescence critique sont comprises dans les $\boldsymbol{k}_c$.

Du fait de la périodicité de $\Delta(\boldsymbol{k})$, tous les points déduits d'un vecteur critique $\boldsymbol{k}_c$ par une translation du réseau réciproque sont aussi les extrémités de vecteurs critiques. Mais le comportement de $I(\boldsymbol{k})$ au voisinage de ces différents points n'est pas le même puisque les facteurs $\exp[i\boldsymbol{k}(\boldsymbol{r}_i - \boldsymbol{r}_j)]$ n'ont pas en général la périodicité du réseau réciproque. On devra donc étudier ce comportement séparément sur tous les points $\boldsymbol{k}_c$ non équivalents, et vérifier en particulier que le numérateur de $I(\boldsymbol{k}_c)$ dans (20) ne s'annule pas en même temps que Δ.

La question centrale est alors la suivante : est-il possible de développer $1/I(\boldsymbol{k})$ au voisinage des vecteurs critiques $\boldsymbol{k}_c$, en faisant apparaitre seulement deux paramètres r_1, et κ_1, comme nous l'avons fait au voisinage de $k = 0$ et des noeuds équivalents.

La réponse est que ceci est en général impossible même pour un cristal cubique, sauf si les conditions suivantes sont réalisées :

(*a*) on s'intéresse uniquement au voisinage immédiat de T_c,

(*b*) le vecteur critique envisagé $\boldsymbol{k}_c$, coïncide avec un noeud $\boldsymbol{\tau}$ du réseau réciproque,

(*c*) le numérateur de (20) n'est pas nul pour $\boldsymbol{k} = \boldsymbol{k}_c$, $T = T_c$. On peut alors développer $1/I(\boldsymbol{k})$ en considérant simultanément $\boldsymbol{k} - \boldsymbol{\tau}$ et $T - T_c$ comme des quantités petites, et on obtient

$$\frac{1}{I(\boldsymbol{k})} = \frac{\sum_i S_i(S_i+1)}{3k_BT\sum_{ij}\tilde{A}_{ji}(\boldsymbol{\tau})\exp\left[i\boldsymbol{\tau}(\boldsymbol{r}_j-\boldsymbol{r}_i)\right]} \times$$

$$\times\left[(T-T_c\frac{\partial\Delta}{\partial T}+\frac{1}{2}\sum_{\epsilon\eta}(k_\epsilon-\tau_\epsilon)(k_\eta-\tau_\eta)\frac{\partial^2\Delta}{\partial k_\epsilon\partial k_\eta}\right] \tag{23}$$

Les dérivées sont prises pour $\boldsymbol{k}=\boldsymbol{\tau}$, $T=T_c$. Le terme d'ordre 0 s'annule puisque $\boldsymbol{\tau}$ est un vecteur critique. Les termes linéaires en $\boldsymbol{k}-\boldsymbol{\tau}$ sont nuls car $\Delta(\boldsymbol{k})$ a la périodicité complète du réseau réciproque et se comporte au voisinage de $\boldsymbol{\tau}$ comme au voisinage de 0. Pour un cristal cubique, on peut alors mettre $1/I(\boldsymbol{k})$ sous la forme

$$\frac{1}{I(\boldsymbol{k})} = r_1^2(\tau)[\kappa_1^2+|\boldsymbol{k}-\boldsymbol{\tau}|^2] \tag{24}$$

avec :

$$\kappa_1^2 = \frac{\sum_i \chi_0^i}{r_1^2(0)\sum_{ij}\chi_{ij}(0)} \tag{22}$$

$$r_1^2(\tau) = \frac{\sum_i S_i(S_i+1)}{18k_BT\sum_{ij}\tilde{A}_{ij}(\boldsymbol{\tau})\exp\left[i\boldsymbol{\tau}\cdot(\boldsymbol{r}_i-\boldsymbol{r}_j)\right]} \times$$

$$\times\sum_\epsilon\left(\frac{\partial^2\Delta}{\partial k_\epsilon^2}\right)_{\substack{T=T_c\\ \boldsymbol{k}=0}} \tag{25}$$

κ_1^2 a la même valeur dans toutes les directions d'opalescence critique. Si l'on exclut le cas des antiferromagnétiques, que nous avons traité séparément,[15] $\boldsymbol{k}=0$ est une telle direction, et l'on peut calculer κ_1 par les formules (22) et (25). Lorsque les hypothèses énumérées ci-dessus sont satisfaites, on peut d'ailleurs montrer que les corrélations entre spins éloignés ont la forme

$$\langle \boldsymbol{S}_{R_i}\,\boldsymbol{S}_{R_j}\rangle = \text{constante}\,\frac{\exp(-\kappa_1|\boldsymbol{R}_i-\boldsymbol{R}_j|)}{|\boldsymbol{R}_i-\boldsymbol{R}_j|} \tag{26}$$

(le signe de la constante dépendant en général du couple (ij) envisagé).

Nous avons dû limiter la définition de $r_1(\boldsymbol{\tau})$ au voisinage immédiat de T_c, car, si $\boldsymbol{\tau}$ n'est pas équivalent à 0, $I(\boldsymbol{\tau})$ n'est plus en général un extremum pour les températures supérieures à T_c. (Le développement (23) contient des termes en $(T-T_c)$ $(k_\epsilon-\tau_\epsilon)$).

Dans la magnétite les conditions d'existence de $r_1(\tau)$ et κ_1 sont satisfaites (voir section 6). Nous avons montré par ailleurs[15] qu'elles ne le sont pas dans les antiferromagnétiques du type MnO.

Les propriétés qualitatives de $I(\boldsymbol{k})$ que nous venons de citer résultent de certaines symétries dans les positions et dans les interactions des ions : quoique déduites d'une approximation de champ moléculaire, elles se retrouveraient dans la solution exacte du problème statistique.

En effet, si l'on fait choix d'une approximation plus raffinée que le champ moleculaire, l'équation (15) est remplacée par une formule plus generale exprimant le champ H_{R_i} comme une combinaison linéaire des moments $\mu_{p,j}$ portes par les sites voisins. Ce type de developpement a été utilisé pour le premiere fois par Yvon.* Les coefficients sont plus compliqués que ceux de (15), mais (*a*) ils decroissent rapidement avec la distance $|\boldsymbol{R}_i-\boldsymbol{R}_j|$ et (*b*) ils on les mêmes proprietes de symetrie que ceux qui figurent dans (15).

4. INTENSITÉ DES FLUCTUATIONS POUR $T < T_c$

Aux températures inférieures à T_c l'approximation quasi-statique de l'équation (6) devient mauvaise car les transferts d'énergie entre le neutron et les spins sont forts (voir Fig. 1 de la réf. 17) et donnent un spectre approximativement discret. Les fluctuations d'aimantation que nous calculons ici ne sont donc plus directement mesurées par les expériences de diffusion des neutrons. Leur calcul est néanmoins instructif car il relie continûment les fluctuations du domaine critique aux fluctuations de basse température, dues aux ondes de spin.

Nous nous limitons toujours à des spins classiques (ce qui est ici une restriction plus sérieuse que pour $T > T_c$); nous désignons par $\bar{\boldsymbol{\mu}}_i$ l'aimantation moyenne de l'ion $\boldsymbol{R}_i$, et par sa variation due aux champs extérieurs $\boldsymbol{H}_{R_i}$. L'équation de champ moléculaire qui remplace (15) s'écrit alors :

$$\boldsymbol{\mu}_{R_i} \equiv \bar{\boldsymbol{\mu}}_i + \sigma\boldsymbol{\mu}_{R_i} = \boldsymbol{\mu}_{R_i}B_i(|\boldsymbol{F}_{R_i}+\boldsymbol{H}_{R_i}|) \tag{27}$$

* Yvon, J., *Cahiers Phys.* A, 28 (1945); B, 31 (1948).

$B_i(H)$ est proportionnelle à la fonction de Brillouin du site i :

$$B_i(H) = \tfrac{1}{2}g\mu_B\{(2S_i+1)\times$$
$$\times \coth \beta g\mu_B(S_i+\tfrac{1}{2})H - \coth \tfrac{1}{2}\beta g\mu_B H\} \qquad (28)$$

$$\boldsymbol{u}_{R_i} = \frac{\boldsymbol{F}_{R_i}+\boldsymbol{H}_{R_i}}{|\boldsymbol{F}_{R_i}+\boldsymbol{H}_{R_i}|} \qquad \boldsymbol{u}_{R^i}^2 = 1$$

$\boldsymbol{F}_{R_i}$ est le champ moleculaire local du site R_i, donné par

$$\boldsymbol{F}_{R^i} = \frac{1}{g^2\mu_B^2}\sum_{jR_j} 2\, J_{R_iR_j}\, \boldsymbol{\mu}_{R_i} \qquad (29)$$

Nous excluons explicitement les configurations obliques pour les aimantations des ions réseaux en l'absence de champ extérieru. Tous les $\bar{\boldsymbol{\mu}}_i$ ont une même orientation ($\pm Oz$.) Dans ces conditions, l'équation vectorielle (27) donne pour les composantes parallèles à Oz ("longitudinales") :

$$\delta\mu^z_{R_i} = \left(\frac{\partial B_i}{\partial H}\right)_{\bar{F}_i} \times$$
$$\times\left[\delta H^z_{R_i} + \frac{1}{g^2\mu_B^2}\sum_{j,R_j} 2\, J_{R_iR_j}\delta\mu^z_{R_j}\right] \qquad (30)$$

($\bar{F}_i$ représente le champ moléculaire sur les sites (i) en l'absence de champ extérieru).

On obtient aussi pour les composantes "transversales" (Ox ou Oy)

$$\delta\mu^x_{R_i} = \frac{\bar{\mu}_i}{F_i}\times$$
$$\times\left[\delta H^x_{R_i} + \sum_{j,R_j}\frac{2\,J_{R_iR_j}\,\delta\mu^z_{R_j}}{g^2\mu_B^2}\right] \qquad (31)$$

Ces équations permettent de calculer le tenseur des susceptibilités

$$\chi^{\epsilon\eta}_{ij}(\boldsymbol{k}) = N\delta_{\epsilon\eta}g^2\mu_B^2\,[A^\epsilon(\boldsymbol{k})^{-1}]_{ij} \qquad (32)$$

ou les $[A^\epsilon]^{-1}$ sont les matrices inverses des A^ϵ définies par

$$A^z_{ij}(\boldsymbol{k}) = \delta_{ij}\frac{g^2\mu_B^2}{\left(\dfrac{\partial B_i}{\partial H}\right)_{\bar{F}_i}} -$$
$$-\sum_{j,R_j} 2\, J_{R_iR_j}\exp[i\boldsymbol{k}(\boldsymbol{R}_i-\boldsymbol{R}_j-\boldsymbol{r}_i+\boldsymbol{r}_j)]$$

$$A^x_{ij}(\boldsymbol{k}) = A^y_{ij}(\boldsymbol{k}) = g^2\mu_B^2\frac{\bar{F}_i}{\bar{\mu}_i}\delta_{ij} -$$
$$-\sum_{j,R_j} 2\, J_{R_iR_j}\exp[i\boldsymbol{k}(\boldsymbol{R}_i-\boldsymbol{R}_j-\boldsymbol{r}_i+\boldsymbol{r}_j)] \qquad (33)$$

En reportant (32) dans (14) on obtiendrait la section efficace dans l'approximation quasi-statique, au sujet de laquelle on peut développer une analyse analogue à celle de la section 3. On retrouve en particulier une opalescence critique au voisinage inférieur au point de transition, dans les directions $\boldsymbol{k}_c$ déjà envisagées.

Pour une étude de ces questions sur des cas simples, nous renverrons le lecteur aux références 14 et nous insisterons plus particulièrement ici sur les différences de comportement entre fluctuations longitudinales et transversales.

(1) Les fluctuations *longitudinales* ne sont notables qu'au voisinage du point de transition (aux températures plus basses elles sont très fortement réduites en amplitude du fait de la saturation). On peut encore leur associer, dans le domaine critique, des paramètres κ_1 et $r_1(\boldsymbol{\tau})$, κ_1^2 étant proportionnel à $(T_c - T)$, moyennant certaines restrictions analogues à celles que nous avions formulées dans le domaine paramagnétique.

(2) Les fluctuations transversales ont un comportement tout différent. Le déterminant $\Delta^x(\boldsymbol{k})$ de la matrice $A^x(\boldsymbol{k})$ s'annule pour $\boldsymbol{k} = \boldsymbol{\tau}$ quelle que soit la température $T < T_c$. La quantité κ_1^2 relative à ces fluctuations est identiquement nulle, et la diffusion des neutrons est anormale, au voisinage des raies magnétiques, dans tout le domaine d'ordre. Aux basses températures, cette diffusion anormale résulte de la présence d'excitations collectives (ondes de spin). En effet la formule :

$$\langle T^x_{i\boldsymbol{k}}T^x_{j,-\boldsymbol{k}}\rangle = Nk_BT[A^x(\boldsymbol{k})^{-1}]_{ij} \qquad (34)$$

déduite de (32) et (13), reste valable, même aux basses températures, pour des spins classiques.

Quoique dérivée ici par une méthode de champ moléculaire, c'est en fait une formule rigoureuse dans le domaine des ondes de spin. Dans ce domaine, on peut en effet récrire l'énergie d'échange (1) sous la forme approchée :

$$-\mathscr{H}_0 = -\sum_{\substack{ij\\R_iR_j}} J_{R_iR_j}\Big(\bar{S}_i\bar{S}_j + S^x_{R_i}S^x_{R_j} + S^y_{R_i}S^y_{R_j} -$$

$$-\frac{1}{2}\frac{\bar{S}_j}{\bar{S}_i}[S(^x_{Ri})^2+(S^y_{Ri})^2]- \qquad (35)$$

$$-\frac{1}{2}\frac{\bar{S}_i}{\bar{S}_j}[(S^x_{Rj})^2+(S^y_{Rj})^2]\Big\}$$

en négligeant seulement des termes d'ordre plus élevé en S^x ou S^y. Ceci peut être mis, à une constante additive près, sous la forme :

$$\mathscr{H}_0=\tfrac{1}{2}\sum_{\boldsymbol{k}}\sum_{ij}\{[A^x_0(\boldsymbol{k})]_{ij}T^x_{i\boldsymbol{k}}T^x_{j,-\boldsymbol{k}}+$$

$$+[A^y_0(-\boldsymbol{k})]_{ij}T^y_{i\boldsymbol{k}}T^y_{j,-\boldsymbol{k}}\} \qquad (36)$$

où la somme $\boldsymbol{k}$ est étendue à la première zone de Brillouin du réseau réciproque, et où les matrices A^x_0, A^y_0 sont précisément les matrices définies par (33) et calculées à température nulle. Enfin, dans le domaine des ondes de spin, les variables $T^x_{i\boldsymbol{k}}$, $T^y_{i\boldsymbol{k}}$ peuvent être considérées comme indépendantes, et leur répartition thermique est pondérée par le facteur $\exp(-\beta\mathscr{H}_0)$. Une formule intégrale classique permet alors de rétablir (34) dans la région de basses températures où les matrices A et A_2 coïncident, à condition que les spins puissent être considérés comme classiques ($k_BT\gg\hbar\omega$). Il est bien connu que l'approximation de champ moléculaire, sous sa forme usuelle (macroscopique) donne des resultats incorrects dans le domaine des ondes de spin. Toutefois, sous la forme locale qui est utilisée ici, c'est une méthode beaucoup plus puissante, qui devient rigoureuse aux deux extremités de l'echelle des températures.

Il n'est pas surprenant d'autre part que la formule (34) qui apparaît ainsi comme la généralisation d'un résultat d'ondes de spin, reste une approximation raisonnable même au voisinage de T_c. Dans le langage de DYSON[18] ceci résulte d'une compensation entre les effets des interactions cinématiques et dynamiques entre ondes de spin.

Pour nos spins classiques on peut dire que les limitations géométriques sur la longueur des S^e_i lendent à diminuer l'intensité des fluctuations transversales quand la température croit; par contre l'énergie des différents modes d'onde de spin décroit en gros comme les aimantations spontanées et ceci augmente l'excitation thermique de ces modes.

5. RELAXATION DES FLUCTUATIONS : EFFET DE RALENTISSEMENT CINÉMATIQUE

Nous donnerons d'abord une description qualitative des phénomènes de relaxation, fondée en partie sur des études antérieures[19] et en partie sur les raisonnements plus spécifiques qui vont suivre. Pour un vecteur $\boldsymbol{k}$ fixé les quantités

$$Q_{ij}(\boldsymbol{k},t)=\exp[i\boldsymbol{k}\cdot(\boldsymbol{r}_i-\boldsymbol{r}_j)][\langle T_{i\boldsymbol{k}}(0)T_{j,-\boldsymbol{k}}(t)\rangle- \\ -\langle T_{i\boldsymbol{k}}\rangle\langle T_{j,-\boldsymbol{k}}\rangle] \qquad (37)$$

ont en général, en fonction de t, un comportement complexe que l'on peut décrire qualitativement par :

(*a*) la valeur pour $t=0$, c'est à dire l'intensité des fluctuations, étudiée dans les sections précédentes;

(*b*) un temps de corrélation τ tel que la fonction étudiée soit négligeable pour $t>\tau$. Ces temps τ sont en géneral de l'ordre de $\hbar/2J$ (ou J est une intégrale d'échange typique) dans tout le domaine paramagnétique. Mais, pour certaines positions particulières de $\boldsymbol{k}$, il arrive que le temps de corrélation d'une certaine combinaison linéaire des Q_{ij} devienne très long devant $\hbar/2J$. Ce "freinage" de la relaxation peut être dû à l'existence de certaines constantes uniformes du mouvement qui limitent le domaine d'extension en phase accessible à partir d'une configuration initiale (freinage cinématique).

Il peut aussi être dû au fait que certaines configurations ordonnées sont presque stables au voisinage supérieur du point d'ordre (freinage thermodynamique). Les deux effets peuvent coexister.

Dans cette section, nous considérons uniquement le premier de ces effets, et nous nous limitons en outre au domaine paramagnétique.

(1) Au voisinage des noeuds du réseau réciproque équivalents à 0, le temps de relaxation de la somme

$$\sum Q_{ij}(0,t)=\langle(\sum_i T_{i,0}(0))(\sum_j T_{j,0}(t))\rangle \\ =\langle[\sum_i T_{i,0}(0)]^2\rangle$$

est infini pour des interactions purement d'échange qui conservent le spin total du spécimen. Lorsque $\boldsymbol{k}$ diffère peu de 0, le temps de corrélation est fini mais très supérieur à $\hbar/2J$. On peut alors décrire cette rétrogradation lente de fluctuations de

VOL. 13 1960

grande longueur d'onde comme une diffusion de spin[11, 19] et poser (pour un cristal cubique)

$$\sum_{ij} Q_{ij}(\boldsymbol{k}, t) = \sum_{ij} Q_{ij}(\boldsymbol{k}, 0) \exp(-\Lambda_1 k^2 t) \quad (38)$$

Un tel coefficient de diffusion Λ_t a été mesuré récemment aux températures élevées ($T \gg T_c$) sur MnF_2[20] et aussi, quoique plus difficilement, sur le fer au voisinage de T_c.[5] Dans ce dernier cas, il est d'ailleurs diminué par un effet de freinage thermodynamique prévu par VAN HOVE[11].

(2) On peut se demander si un tel allongement du temps de corrélation se retrouve au voisinage des vecteurs critiques non équivalents à 0.

En fait, l'amplitude de diffusion pour $\boldsymbol{k} = \boldsymbol{\tau}$ est une certaine combinaison linéaire des aimantations des différents sous réseaux; elle peut être décomposée en un terme proportionnel à l'aimantation totale (temps de corrélation infini) et en un terme qui fait intervenir les différences d'aimantation entre sous réseaux : ces différences disparaissent en un temps microscopique de l'ordre de $\hbar/2J$ et donnent lieu à une composante de diffusion inélastique, même lorsque $\boldsymbol{k}$ coïncide exactement avec τ.

Nous allons maintenant discuter cette décomposition plus en détail. Lorsque le vecteur $\boldsymbol{k}$ envisagé coïncide avec un noeud du réseau $\boldsymbol{\tau}$, nous n'avons à étudier que les aimantations globales $M_1 \ldots M_n$ des différents sous réseaux. L'intensité $I(\boldsymbol{k})$, que nous considérerons seule ici, devient par une transformation simple de (20)

$$I(\boldsymbol{\tau}) = \frac{\sum_{ij} \langle M_i M_j \rangle \exp[i\boldsymbol{\tau}(\boldsymbol{r}_i - \boldsymbol{r}_j)]}{k_B T \sum_i \chi_0^i} \quad (39)$$

M_i désigne la composante de M_i selon un axe arbitraire (Ox par exemple) et χ_0^i la susceptibilité du sous réseau i en l'absence d'interactions d'échange. La fraction de $I(\boldsymbol{\tau})$ qui correspond à une diffusion élastique s'écrit d'après (5) (voir réf. 11) :

$$I_{el} = \frac{\sum_{ij} \langle M_i(t=0) M_j(t=\infty) \rangle \exp[i\boldsymbol{\tau}(\boldsymbol{r}_i - \boldsymbol{r}_j)]}{k_B T \sum \chi_0^i} \quad (40)$$

et c'est cette fonction que nous allons déterminer maintenant par un calcul purement thermodynamique. Nous chercherons tout d'abord la variation du potential thermodynamique Φ lorsque le système de spins passe reversiblement de son état le plus probable à un état *voisin*, de même température, où chacune des aimantations M_i a une valeur prescrite.

Dans ces conditions l'orientation moyenne imposée à chaque spin est très inférieure à sa valeur de saturation, et l'entropie augmente d'une quantité

$$s = -\frac{1}{2T} \sum_i \frac{M_i^2}{\chi_0^i} \quad (41)$$

Dans la même transformation, l'énergie d'échange, varie de

$$U = -\tfrac{1}{2} \sum_{ij} \lambda_{ij} M_i M_j \quad (42)$$

avec, dans l'approximation de champ moléculaire

$$\lambda_{ij} = \frac{1}{g^2 \mu_B^2} \sum_{R_j} 2 J_{R_i R_j} = \lambda_{ji} \quad (43)$$

Le potentiel thermodynamique cherché est

$$\Phi = U - Ts - \sum_i M_i H_i \quad (44)$$

où nous avons supposé que des champs extérieurs. H_i différents étaient appliqués sur les sous réseaux ($i = 1, \ldots, n$). On vérifiera aisément sur l'expression ci-dessus que le minimum de Φ correspond aux valeurs des aimantations prévues par l'approximation de champ moléculaire.

Dans la solution exacte dy modèle statistique, on retrouverait un potentiel de même forme, les coefficients λ_{ij} ayant seulement des valeurs un peu modifiées.

La connaissance de Φ détermine la probabilité d'une configuration ($M_1 \ldots M_n$), qui est proportionnelle à $\exp[-\beta \Phi d(M_1 \ldots dM_n)]$ Ceci permet de rétablir les formules

$$\langle M_i M_j \rangle = g^2 \mu_B^2 [A^{-1}(0)]_{ij} \quad (45)$$

où les matrices A et A^{-1} sont définies comme dans section 3.

Il est utile d'introduire ici un espace à n dimensions (espace des aimantations) dont les M_i sont les coordonnées orthogonales. Dans l'espace des aimantations les surfaces (Σ) définies par $\Phi =$ constante, sont du type ellipsoïde pour $T > T_c$. Pour $T = T_c$ elles dégénèrente, et l'on peut

définir une direction de facile fluctuation $(D_1, D_2, ..., D_n)$ telle que

$$\Phi_{T_c}(M_1+\xi D_1),\ M_2+\xi D_2, ..., M_n+\xi D_n) \equiv \Phi_{T_c}(M_1, M_2, ..., M_n) \tag{46}$$

Nous supposerons, conformément aux situations usuelles que dans toutes les autres directions Φ se comporte comme une forme définie positive même pour $T = T_c$. (Nous verrons plus loin, à propos d'un exemple plus complexe où cette hypothèse n'est pas satisfaite).

A chaque température T l'examen des surfaces (Σ) permet de situer qualitativement les fluctuations qui sont prédominantes. Nous illustrons leur aspect par la Fig. 1, qui est relative au cas de deux sous réseaux.

Considérons maintenant dans l'espace des aimantations une certaine configuration initiale $(M_1, ..., M_n)$. Du fait des interactions d'échange, cette configuration évolue. L'aimantation totale du spécimen étant une constante du mouvement, le point représentatif se déplace sur un hyperplan $\Sigma M_i =$ constante. Il n'existe pas d'autre constante du mouvement pour limiter ses trajectoires possibles, car l'énergie peut être échangée entre la composante de Fourier particulière que nous étudions ici et toutes les autres composantes. Dans ces conditions il est conforme aux postulats généraux de la thermodynamique d'admettre que le point représentatif tend vers une position d'équilibre P qui correspond au minimum du potentiel thermodynamique dans l'hyperplan auquel il est lié (Fig. 1).

Les coordonnées P_i de P peuvent être calculées d'après cette condition en fonction des coordonnées M_i de la configuration initiale

$$\left.\begin{aligned} \sum_i P_i &= \sum_i M_i \\ \frac{\partial \Phi}{\partial M_1}(P) &= \frac{\partial \Phi}{\partial M_2}(P) = ... = \frac{\partial \Phi}{\partial M_n} P \end{aligned}\right\} \tag{47}$$

Ce système admet une solution unique dans tous les cas, sauf pour les antiferromagnétiques (au point de Néel) que nous excluons ici explicitement.

Elle s'écrit :

$$P_j = \frac{\sum_l [A^{-1}(0)]_{jl}}{\sum_{lm} [A^{-1}(0)]_{lm}} \sum_p M_q \tag{48}$$

Les points M et $P(M)$ correspondent respectivement à l'état initial et à l'état final (après un temps très long) du système de spins, c'est à dire aux quantités $M(t=0)$ et $M(t=\infty)$ de la formule

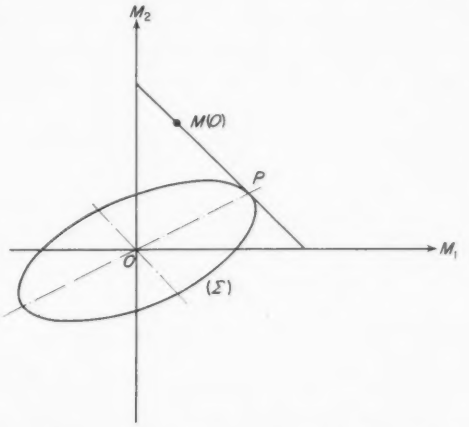


FIG. 1. Diagramme des aimantations dans le cas où il y a 2 atomes par maille. Le point M dont les coordonnées sont les aimantations des deux sous réseaux relaxe sur la droite $M_1+M_2 = C^{te}$ jusqu'au point P de contact avec à l'ellipse (Σ) d'equation $\Phi = C^{te}$.

(40). Nous pouvons récrire celle-ci sous la forme

$$\frac{I_{el}(\tau)}{I(\tau)} = \frac{\sum_{ij} \langle M_i P_j(M) \rangle \exp[i\tau\cdot(r_i - r_j)]}{\sum_{ij} \langle M_i M_j \rangle \exp[i\tau\cdot(r_i - r_j)]} \qquad \frac{\sum_{ijlq} \langle M_i M_j \rangle [A^{-1}(0)]_{jl} \exp[i\tau(r_i - r_j)]}{\{\sum_{ij} \langle M_i M_j \rangle \exp[i\tau(r_i - r_j)]\}\ [\sum_{lm} [A^{-1}(0)]_{lm}]} \tag{49}$$

Cette formule peut être ramenée à des quantités connues en utilisant (45). Finalement, en remarquant que la matrice $A^{-1}(0)$ est symétrique, on obtient

$$\frac{I_{el}(\tau)}{I(\tau)} = \frac{|\sum_{iq} [A^{-1}(0)]_{iq} \exp(i\tau\cdot r_i)|^2}{[\sum_{lm} [A^{-1}(0)]_{lm}][\sum_{iq} [A^{-1}(0)]_{iq} \exp[i\tau(r_i - r_q)]} \tag{50}$$

Nous allons maintenant discuter les valeurs de ce rapport. Il est différent de 1 pour $\tau \neq 0$ et pour les températures $T > T_c$. En particulier, dans la

limite des températures élevées ($T \gg T_c$) il prend la valeur simple

$$\left[\frac{I_{\text{él}}(\boldsymbol{\tau})}{I(\boldsymbol{\tau})}\right]_{T=\infty} = \left|\frac{\sum_i S_i(S_i+1)\exp(i\boldsymbol{\tau}\cdot\boldsymbol{r}_i)}{\sum_i S_i(S_i+1)}\right|^2 \quad (51)$$

Toutefois, lorsque T se rapproche de T_c, le rapport (50) tend vers 1. Ceci peut être démontré de la façon suivante; posons :

$$v_l(\boldsymbol{\tau}) = \sum_m \tilde{A}_{lm}(0)\exp(i\boldsymbol{\tau}\cdot\boldsymbol{r}_m) \quad (52)$$

En multipliant le numérateur et le dénominateur de (50) par $\Delta(0)$ on obtient

$$\frac{I_{\text{él}}(\boldsymbol{\tau})}{I(\boldsymbol{\tau})} = \frac{(\sum_i v_i(0)\exp(i\boldsymbol{\tau}\cdot\boldsymbol{r}_i))\,(\sum_l v_l(\boldsymbol{\tau}))}{(\sum_i v_i(\boldsymbol{\tau})\exp[(i\boldsymbol{\tau}\cdot\boldsymbol{r}_i)]\,(\sum_l v_l(0))} \quad (53)$$

Les quantités v_i satisfont d'autre part aux relations

$$\sum_j A_{ij}(0)v_j(\boldsymbol{\tau}) = \sum_{jl} A_{ij}(0)\tilde{A}_{jl}(0)\exp(-i\boldsymbol{\tau}\cdot\boldsymbol{r}_l)$$
$$= \Delta(0)\exp(-i\boldsymbol{\tau}\cdot\boldsymbol{r}_i)$$

Au point de transition le nombre de droite de cette dernière équation est nul. Pour chaque $\boldsymbol{\tau}$, les n quantités $v_j(\boldsymbol{\tau})$ forment donc un vecteur propre de la matrice A correspondant à la valeur propre 0 (donc parallèle à la direction de facile fluctuation D que nous avons supposée unique). Ceci peut être écrit.

$$v_i(\boldsymbol{\tau}) = \xi(\boldsymbol{\tau})v_i(0)$$

(où ξ est indépendant de i) et montre d'après (53) que

$$\left[\frac{I_{\text{él}}(\boldsymbol{\tau})}{I(\boldsymbol{\tau})}\right]_{T=T_c} = 1 \quad (54)$$

Ce résultat apparait clairement dans l'espace des aimantations. Le vecteur $\boldsymbol{OM}_i$ a une composant $\boldsymbol{PM}$ à relaxation rapide et une $\boldsymbol{OP}$ à relaxation bloquée. Pour $T = T_c$ le point P est situé sur la direction de facile fluctuation D. Les fluctuations de $\boldsymbol{OP}$ deviennent infinies, alors que celles de $\boldsymbol{PM}$ restent finies. La composante inélastique de la diffusion reste finie à la température critique, mais le rapport de cette composante à l'intensité totale tend vers 0.

Les conclusions auxquelles nous sommes parvenus peuvent être résumées par la formule

$$\sum_{ij}\frac{1}{2\pi}\int_{-\infty}^{\infty} dt\,\exp(-i\omega t)Q_{ij}(\boldsymbol{\tau},t)$$
$$= \frac{\sum_{ij}Q_{ij}(\boldsymbol{\tau},0)}{I(\boldsymbol{\tau})}[I_{\text{él}}(\boldsymbol{\tau})\delta(\omega)+(I(\boldsymbol{\tau})-I_{\text{él}}(\boldsymbol{\tau}))f_\tau(\omega)]$$

$f_\tau(\omega)$ est une fonction régulière de ω, normalisée ($\int f_\tau(\omega)d\omega = 1$) et telle que les transferts d'énergie $\hbar\omega$ correspondants soient de l'ordre de grandeur des intégrales d'échange.

Le problème qui se présente maintenant est d'étendre ces résultats au cas où $\boldsymbol{k}$ ne coïncide plus exactement avec $\boldsymbol{\tau}$ (mais en reste très voisin).

Il est encore possible, dans ces conditions, de distinguer deux composantes dans le spectre des fréquences ω dont le rapport d'intensité est encore donné par (50) si l'on s'en tient au premier ordre significatif en $\boldsymbol{k} - \boldsymbol{\tau}$.

L'une des composantes provient d'une relaxation avec transfert d'aimantation entre sous réseaux qui s'effectue en un temps microscopique : son spectre $f_{\boldsymbol{k}}(\omega)$ ne subit pas de modification remarquable. L'autre composante correspond à une égalisation lente de l'aimantation sur des distances $\sim|\boldsymbol{k}-\boldsymbol{\tau}|^{-1}$ grandes par rapport à la maille, et peut être décrite par le coefficient de diffusion Λ_1, déjà introduit. Au total :

$$\sum_{ij}\frac{1}{2\pi}\int_{-\infty}^{\infty} dt\,\exp(-i\omega t)Q_{ij}(\boldsymbol{k},t)$$
$$= \frac{\sum_{ij}Q_{ij}(\boldsymbol{k},0)}{I(\boldsymbol{\tau})}\left[\frac{1}{\pi}I_{\text{él}}(\boldsymbol{\tau})\frac{\Lambda_1(\boldsymbol{k}-\boldsymbol{\tau})^2}{\omega^2+\Lambda_1|(\boldsymbol{k}-\boldsymbol{\tau})^4}+\right.$$
$$\left.+(I(\boldsymbol{\tau})-I_{\text{él}}(\boldsymbol{\tau}))f_{\boldsymbol{k}}(\omega)\right] \quad (55)$$

La formule (55) est valable à toutes les températures $T \geqslant T_c$. Elle suppose $|\boldsymbol{k}-\boldsymbol{\tau}|a_0 \ll 1$ où a_0 est une dimension typique de la maille. Elle fait apparaitre un terme quasi élastique dont l'intensité est très grande au voisinage du point de transition, et un terme inélastique dont l'intensité reste toujours finie. Enfin, il faut observer que (55) ne s'applique pas aux antiferromagnétiques, sur lesquels nous reviendrons d'ailleurs dans la section 6.

VOL. 13 1960

6. RELAXATION DES FLUCTUATIONS : EFFET DE RALENTISSEMENT THERMODYNAMIQUE

Nous allons chercher à voir ici si la thermodynamique des phénomènes irréversibles permet de faire des prédictions sur la relaxation de certaines fluctuations. En premier lieu nous calculerons la source d'entropie $\dot{\sigma}$ lors d'une évolution isotherme, à champs extérieurs constants, l'état du système de spins étant spécifié par les aimantations globales M_i des différents sous réseaux. (Nous nous plaçons toujours au-dessus de T_c). Pour une transformation infinitésimale de ce type l'energie interne V et le travail W fourni aux spins par les champs extérieurs varient respectivement de :

$$dV = -\sum_{ij} \lambda_{ij} M_i \, dM_j$$

$$dW = \sum_i H_i \, dM_i$$

La chaleur engendrée est $dQ = dV - dW$ et la source d'entropie est donnée par

$$\begin{aligned} T\dot{\sigma} &= \frac{dQ}{dt} - T\frac{ds}{dt} \\ &= \sum_i \frac{\partial \Phi}{\partial M_i} \frac{dM_i}{dt} \end{aligned} \tag{56}$$

Nous sommes donc en droit de définir un système de forces $X_i = -\partial\Phi/\partial M_i$ et un système de flux $\dot{M}_i = \partial M_i/\partial t$. Moyennant certaines hypothèses très restrictives sur lesquelles nous reviendrons, il sera *parfois* possible de postuler un système de relations linéaires entre les flux et les forces

$$\dot{M}_i = \sum_i \Gamma_{ij} X_i \tag{57}$$

$$= -\sum_l R_{il} M_l \tag{58}$$

$$R_{il} = \frac{\Gamma_{il}}{\chi_0^i} - \Sigma \Gamma_{ij} \lambda_{jl} \tag{59}$$

On peut illustrer la signification de la matrice Γ en imaginant l'expérience (fictive) suivante : à tous les temps $t < 0$ un système de champs H_i est appliqué aux différents sous réseaux. Au temps 0 ces champs sont brusquement annulés. Les aimantations se mettent à varier, et leurs dérivées initiales sont

$$(\dot{M}_i)_{t=0} = -\sum_j \Gamma_{ij} H_j \tag{60}$$

Il est naturel de supposer que toutes ces dérivées sont finies et que par conséquent les Γ_{ij} sont finis à toutes températures $T \geqslant T_c$. D'autre part, l'hamiltonien d'échange étant invariant par renversement du sens du temps, la matrice Γ (en l'absence de champ extérieur) est symétrique ($\Gamma_{ij} = \Gamma_{ji}$) d'apres le théorème d'Onsager. (La matrice R, par contre, n'est pas symétrique en géneral). Enfin la constance du spin total impose les relations

$$\begin{aligned} \sum_i \dot{M}_i &= -\sum_i R_{il} M_l \equiv 0. \\ \sum_i R_{il} &= 0. \end{aligned} \tag{61}$$

Nous appliquons, à titre d'exemple, ces considérations formelles aux *antiferromagnétiques à deux sous réseaux* (du type cubique simple ou cubique centré), et nous verrons ensuite quel est leur domaine de validité dans ce cas.

Le potentiel thermodynamique s'écrit ici

$$\begin{aligned} \Phi = \frac{1}{2\chi_0}(M_1^2+M_2^2) - \tfrac{1}{2}\lambda(M_1^2+M_2^2) - \\ - \lambda' M_1 M_2 - H_1 M_1 - H_2 M_2 \end{aligned} \tag{62}$$

$$\lambda = \lambda_{11} = \lambda_{22} \qquad \lambda' = \lambda_{12} < 0$$

(χ_0 est la susceptibilité d'un sous-réseau isolé). Les équations d'évolution (en l'absence de champs extérieurs) déduites de (57) et (59) s'écrivent

$$\begin{aligned} \dot{M}_1 &= \Gamma\left[\frac{1}{\chi_0} - \lambda + \lambda'\right](M_2 - M_1) \\ \dot{M}_1 + \dot{M}_2 &= 0 \end{aligned} \tag{63}$$

où l'on a fait usage de la symétrie des Γ_{ij}, des relations (61) et où l'on a posé $\Gamma_{11} = \Gamma$.

La quantité intéressante sur les raies de surstructure est $M_1 - M_2$ qui satisfait à

$$\frac{d}{dt}\Big(M_1 - M_2\Big) = \frac{1}{\tau_0}\Big(M_2 - M_1\Big) \tag{64}$$

avec

$$\frac{1}{\tau_0} = \frac{2\Gamma}{\chi_0}\Big(1 - \lambda\chi_0 + \lambda'\chi_0\Big) \tag{65}$$

Le membre de droite de (65) s'annule précisément au point de Néel, comme il est possible de le voir en définissant ce point à partir de (62). τ_0 est donc

infini à cette température, et très long aux températures voisines. Nous avons là un cas typique de "freinage thermodynamique".

L'évolution irréversible de $M_1 - M_2$ à partir d'une valeur initiale prescrite est lente par rapport à l'échelle microscopique du temps $\hbar/2J$. Lorsque l'on s'intéresse à des temps très supérieurs à $\hbar/2J$ il est alors plausible d'admettre la validité de (57). D'autre part, la direction $M_1 = -M_2$ est la direction de facile fluctuation. Les fluctuations spontanées de $M_1 - M_2$ sont quasi-macroscopiques au voisinage du point de Néel. Il est correct dans ces conditions, de supposer que la rétrogradation des fluctuations spontanées à même dépendance temporelle qu'une rétrogradation irréversible décrite par (64). Ceci permet d'écrire :

$$\langle (M_1-M_2)_0\,(M_1-M_2)_t\rangle = \langle (M_1-M_2)^2\rangle \exp(-t/\tau_0) \qquad (66)$$

Cette formule est donc valable pour $(T - T_c)/T_c \ll 1$. (T_c est le point de Néel). Elle montre que le spectre en énergie de la section efficace (5), lorsque $\boldsymbol{k}$ coïncide avec une raie de surstructure, a la forme de Lorentz, que les transferts d'énergie sont de l'ordre de $\hbar/\tau_0$ et qu'ils s'annulent au point de Néel d'après (65).

Si l'on attribue à λ et λ' leurs valeurs dans l'approximation de champ moléculaire (indépendantes de la température), on peut d'ailleurs donner à (66) la forme détaillée

$$\langle (M_1-M_2)_{t=0}(M_1-M_2)_t\rangle = \frac{2NS(S+1)}{3}\,\frac{T}{T-T_c}\exp\left\{-\frac{2\Gamma(T-T_c)t}{\chi_0 T}\right\} \qquad (67)$$

où N est le nombre de mailles magnétiques.

Un cas plus complexe mais également intéressant du point de vue expérimental est celui des antiferromagnétiques du type MnO. Nous ne donnons ici qu'un résumé de l'étude correspondante pour la direction (111) qui est la plus importante du point de vue de l'opalescence critique. On doit distinguer pour cette structure cubique à faces centrées, huit sous réseaux $(1_+2_+3_+4_+$ et $1_-2_-3_-4_-$. Chaque site (1_+) par exemple est couplé à ses douze premiers voisins $(2\times(2_+)+2\times(3_+)+2\times(4_+)+2\times(2_-)+2\times(3_-)+2\times(4_-))$ et aussi, antiferromagnétiquement, à ses six deuxième voisins (1_-). L'amplitude de diffusion est proportionnelle à

$$a_{111}(M) = M_{1+}-M_{1-}+M_{2-}-M_{2+}+M_{3-}-M_{3+}+ + M_{4-}-M_{4+} \qquad (68)$$

L'espace des aimantations a 8 dimensions. Pour $T > T_c$ les coordonnées du point $P(M)$ défini par (47) sont égales

$$P_{i+}(M) = \tfrac{1}{8}\sum_1^4 (M_{i+}+M_{i-}) \qquad (69)$$

On voit immédiatement que $a_{111}\,(P)$ est nul. L'analyse de la section 5 montre alors qu'il n'y a pas de composante strictement élastique à la diffusion sur la raie (111) :

$$\langle a_{111} = (t=0)a_{111}(t=\infty)\rangle = 0 \qquad (70)$$

Examinons maintenant les effets de freinage thermodynamique. Dans les cas où il y a effectivement opalescence critique sur la raie (111)[15] on ne trouve pas, comme d'habitude, une direction de facile fluctuation au point de Néel, mais une infinité de telles directions, qui sont contenues dans un sous espace ϵ à 4 dimensions de l'espace des aimantations. ϵ est défini par les équations :

$$M_{i+}+M_{i-} = 0 \qquad (71)$$

Si (Q) est la projection du "vecteur" (M) dans ϵ on peut montrer que $a_{111}\,(Q) = a_{111}\,(M)$: il suffit donc d'étudier les relaxations de (M) dans le sous espace ϵ. Ici encore, cette relaxation est *lente* au voisinage de T_c car pour $T = T_c$ le potentiel Φ est constant dans ϵ et les forces X_i définies plus haut s'annulent. D'autre part les propriétés de symétrie du potentiel dans ϵ font que la relaxation peut être décrite par une seule exponentielle décroissante. Finalement on obtient (pour $T-T_c/T_c \ll 1$) :

$$\langle a_{111}(0)a_{111}(t)\rangle = \langle a_{111}^2\rangle \exp(-t/\tau_0) \qquad (72)$$

τ_0 étant long devant l'échelle microscopique des temps et infini pour $T = T_c$.

Ces propriétés remarquables pourraient être mises en évidence par des expériences de diffusion inélastique des neutrons.

Mentionnons pour terminer un autre exemple, déjà connu, de freinage thermodynamique : il s'agit de l'annulation du point de Curie du coefficient de diffusion de spin Λ_1, dans les ferromagnétiques. Cette annulation se retrouve dans

VOL. 13 1960

les ferrimagnétiques, comme on peut le vérifier par la méthode originale de VAN HOVE[11].

Pour toutes ces substances la relaxation des fluctuations de grande longueur d'onde est donc freinée simultanément par l'effet cinématique et par l'effet thermodynamique. Pour les antiferromagnétiques au contraire, l'effet thermodynamique joue seul dans les directions d'opalescence critique.

7. APPLICATION À LA MAGNÉTITE

Le cas de la magnétite constitue une très bonne illustration des formules générales établies plus haut. Le groupe de translations est celui du cube à faces centrées. Il y a six ions magnétiques dans la maille, de coordonnées

$$\left.\begin{aligned} \boldsymbol{r}_1 &= \begin{pmatrix}0\\0\\0\end{pmatrix} \quad \boldsymbol{r}_2 = \begin{pmatrix}\frac{1}{4}\\\frac{1}{4}\\\frac{1}{4}\end{pmatrix} a \quad \boldsymbol{r}_3 = \begin{pmatrix}\frac{5}{8}\\\frac{5}{8}\\\frac{5}{8}\end{pmatrix} a; \\ \boldsymbol{r}_4 &= \begin{pmatrix}\frac{5}{8}\\\frac{3}{8}\\\frac{3}{8}\end{pmatrix} a; \quad \boldsymbol{r}_5 = \begin{pmatrix}\frac{3}{8}\\\frac{5}{8}\\\frac{3}{8}\end{pmatrix} a; \quad \boldsymbol{r}_6 = \begin{pmatrix}\frac{3}{8}\\\frac{3}{8}\\\frac{5}{8}\end{pmatrix} a \end{aligned}\right\} \tag{73}$$

où a est la longueur de l'arête du cube. Les ions des sites (1) et (2) sont des ions ferriques et portent un spin $S_a = 5/2$. Les ions des sites (3), (4), (5) et (6) sont pour moitié des ions ferriques et pour moitié des ions ferreux. Aux températures qui nous intéressent la répartition de ces deux types d'ions est totalement aléatoire, puisque l'ordre de position est détruit à partir d'une température beaucoup plus basse ($T = 118°$ K). Le fond continu dans la diffusion magnétique des neutrons qui résulte de ce désordre de position est très faible (et sans anomalie de répartition aux températures comparables à T_2).

Nous le négligerons dans la suite, et nous définirons un "spin moyen" pour les sites (3), (4), (5) ou (6) (dans le domaine paramagétique) par la relation

$$S_d(S_d+1) = 2\left(6+\frac{35}{4}\right) = \frac{59}{8}$$

Nous supposons, conformément aux résultats du calcul initial de NÉEL (22) que deux interactions seulement ne sont pas négligeables : celles (J_{aa}) qui couplent entre eux les ions a premiers voisins, et celles (J_{ad}) qui couplent les ions α à leurs voisins d. Pour les calculs numériques nous avons pris

$$\frac{J_{aa}}{k_B} = -18° \text{ K} \qquad \frac{J_{ad}}{k_B} = -23° \text{ K}$$

valeurs qui reproduisent assez bien la susceptibilité paramagnétique expérimentale. Enfin une dernière approximation qui s'impose dans l'état actuel de nos connaissances expérimentales consiste à attribuer le même facteur de forme à tous les ions ferriques ou ferreux. La section efficace est alors, dans le domaine paramagnétique, proportionnelle à la quantité $I(\boldsymbol{k})$ définie par (20).

Pour le calcul pratique nous introduirons la matrice

$$A'_{ij}(\boldsymbol{k}) = A_{ij}(\boldsymbol{k}) \exp[i\boldsymbol{k}(\boldsymbol{r}_i - \boldsymbol{r}_j)] \tag{74}$$

dont l'expression détaillée est la suivante :

$$\frac{(A')}{2J_{ad}} = -\begin{Vmatrix} T/\alpha & \alpha\Sigma_{12} & \Sigma_{13} & \Sigma_{14} & \Sigma_{15} & \Sigma_{16} \\ \alpha\Sigma_{12}^* & T/a & \Sigma_{13}^* & \Sigma_{14}^* & \Sigma_{15}^* & \Sigma_{16}^* \\ \Sigma_{13}^* & \Sigma_{13} & T/b & 0 & 0 & 0 \\ \Sigma_{14}^* & \Sigma_{14} & 0 & T/b & 0 & 0 \\ \Sigma_{15}^* & \Sigma_{15} & 0 & 0 & T/b & 0 \\ \Sigma_{16}^* & \Sigma_{16} & 0 & 0 & 0 & T/b \end{Vmatrix} \tag{75}$$

Les notations nouvelles introduites sont

$$a = -\frac{2}{3k_B} J_{ad}\, S_a(S_a+1)$$

$$b = -\frac{2}{3k_B} J_{ad}\, S_d(S_d+1)$$

$$\alpha = J_{aa}/J_{ad}$$

$$\Sigma_{12} = 4(\cos\pi h/2 \ \cos\pi k/2 \ \cos\pi l/2 - i \ \sin \pi h/2 \ \sin\pi k/2 \ \sin\pi l/2)$$

$$\Sigma_{13} = \exp[i\pi(h+k+l/4)]\,[\exp(-i\pi h) + \exp(-i\pi k) + \exp(-i\pi l)]$$

$$\Sigma_{14} = \exp[i\pi(h-k-l/4)]\,[\exp(-i\pi h) + \exp(i\pi k) + \exp(i\pi l)]$$

$$\Sigma_{15} = \exp[i\pi(-h+k-l/4)]\,[\exp(i\pi h) + \exp(-i\pi k) + \exp(i\pi l)]$$

$$\Sigma_{16} = \exp[i\pi(-h-k+l/4)]\,[\exp(i\pi h) + \exp(i\pi k) + \exp(-i\pi l)]$$

$2\pi h$, $2\pi k$, $2\pi l$ sont les composantes du vecteur $a\,\boldsymbol{k}$.

Il sera utile de poser

$$\sigma = |\Sigma_{13}|^2+|\Sigma_{14}|^2+|\Sigma_{15}|^2+|\Sigma_{16}|^2$$
$$\rho = \Sigma_{13}^2+\Sigma_{14}^2+\Sigma_{15}^2+\Sigma_{16}^2$$
$$\zeta = \Sigma_{13}+\Sigma_{14}+\Sigma_{15}+\Sigma_{16}$$

On peut alors récrire la formule (20) sous la forme

$$I(\boldsymbol{k}) = \frac{3k_BT}{\sum_{ij} S_i(S_i+1)} \sum_{ij} [A'^{-1}(\boldsymbol{k})]_{ij}$$
$$= \frac{1}{2+4\frac{b}{a}} \frac{g(\boldsymbol{k},T)}{G(\boldsymbol{k},T)} \qquad (76)$$

où

$$g\boldsymbol{k}(,T) \equiv \left(2+4\frac{b}{a}\right) T^4+(\gamma_1 b+\gamma_2\alpha a)T^3 +$$
$$+ (\gamma_3 b^2+\gamma_4 ab)T^2+ \gamma_5 ab^2T+\gamma_6 ab^3$$

$$G(\boldsymbol{k},T) \equiv T^4+(\Gamma_1 ab+\Gamma_2\alpha^2a^2)T^2 +$$
$$+\Gamma_3\alpha a^2bT+\Gamma_4 a^2b^2$$

avec

$$\gamma_1 = -2(\zeta+\zeta^*) \qquad \gamma_2 = -(\Sigma_{12}+\Sigma_{12}^*)$$
$$\gamma_3 = -2(40-\zeta\zeta^*)$$
$$\gamma_4 = 2\alpha\zeta\Sigma_{12}^*+2\alpha\zeta^*\Sigma_{12}+\rho+\rho^*-2\sigma-4\sigma^2|\Sigma_{12}|^2$$
$$\gamma_5 = \alpha\Sigma_{12}(4\rho^*-\zeta^{*2})+\alpha(4\rho-\zeta^2)+2(\sigma-\rho)\zeta^* +$$
$$+ 2(\sigma-\rho^*)\zeta$$
$$\gamma_6 = 4\sigma^2-4\rho\rho^*-2\sigma\zeta\zeta^*+\rho\zeta^{*2}+\rho^*\zeta^2$$
$$\Gamma_1 = -20 \qquad \Gamma_2 = -|\Sigma_{12}|^2$$
$$\Gamma_3 = \rho\Sigma_{12}^*+\rho^*\Sigma_{12} \qquad \Gamma_4 = \sigma^2-\rho\rho^*$$

$G(\boldsymbol{k},T)$ est une fonction de $\boldsymbol{k}$ qui a la période du réseau réciproque, qui est positive pour $T > T_c$ et qui s'annule pour $T = T_c$ sur les noeuds $\boldsymbol{k} = \boldsymbol{\tau}$ du réseau réciproque. $g(\boldsymbol{k},T)$ a en revanche une période plus grande dont les translations élémentaires engendrent une structure cubique à face centrée dans le réseau réciproque.

Pour déterminer les paramètres de l'opalescence critique sur les différents noeuds, on développe G au voisinage de $\boldsymbol{k} = \boldsymbol{\tau}$ et $T = T_c$ comme nous

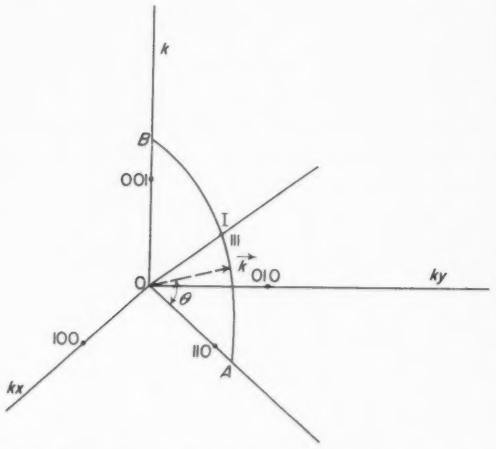


FIG. 2. Définition de l'angle θ utilisé dans les Figs. 4, 6 et 7. Le vecteur $\boldsymbol{k}$ se déplace sur le segment OI et sur le quart de cercle $AIB-I$ est le noeud (111).

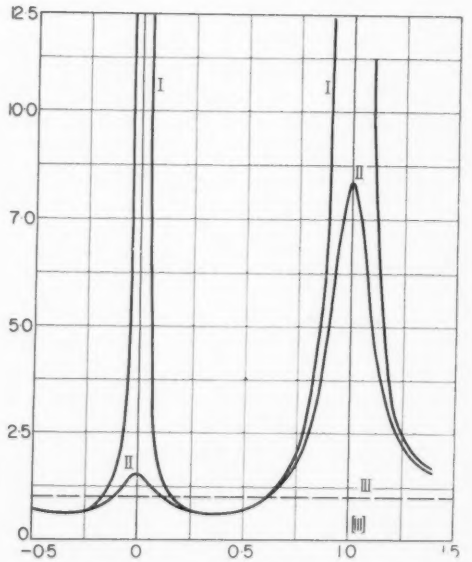


FIG. 3. Variations de I pour Fe_3O_4 quand $\boldsymbol{k}$ décrit le segment OI. On a posé $k_x = k_y = k_z = 2\pi\lambda/\alpha$ et on a tracé la courbe pour plusieurs valeurs de la température T.

Courbe I : $T = T_c$
Courbe II : $T = T_c+40^0$
Courbe III: $T = \infty$.

l'avons expliqué dans la section 3, et on obtient

$$\left.\begin{aligned}\kappa_1^2 &= 0{,}494\,\frac{T-T_c}{T_c}\ \text{Å}^{-2}\\ r_1^2 &= 10^9\,\frac{7220}{g(\tau,T_c)}\ \text{Å}^2\end{aligned}\right\} \qquad (77)$$

Les valeurs des fonctions g qui figurent dans cette dernière expression sont pour $T = T_c$:

$g(000) = 187 \,.\, 10^9$ $\quad$ $g(200) = 0$

$g(220) = 550 \,.\, 10^9$ $\quad$ $g(400) = 3656 \,.\, 10^9$

$g(111) = 1230 \,.\, 10^9$ $\quad$ $g(333) = 4 \,.\, 10^9$

$g(222) = 1375 \,.\, 10^9$

On voit que l'opalescence critique est forte autour des raies (111) et (222), mais faible sur les raies (000) et (333).

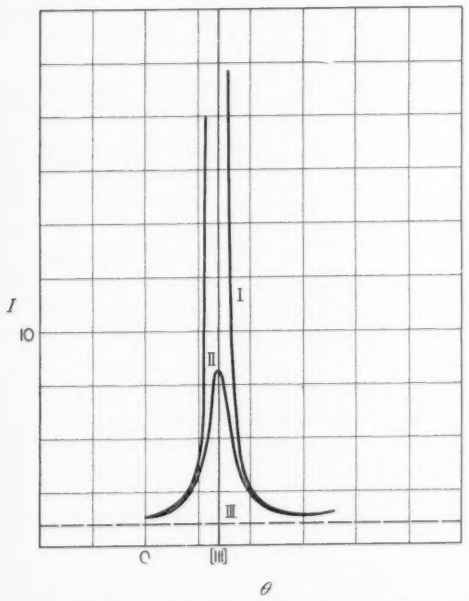


FIG. 4. Variations de I quand $\boldsymbol{k}$ décrit le quart de cercle AIB. θ est l'abscisse curviligne comptée depuis le plan (001). Les différentes températures sont dénotées comme sur la Fig. 3.

Les Figs. 3 et 4 donnent la variation de $I(\boldsymbol{k})$ lorsque $\boldsymbol{k}$ varie :

(*a*) sur l'axe 111,

(*b*) sur un cercle passant par le point (111) et contenu dans le plan des directions (111) et (001) (voir Fig. 2).

Les Figs. 5, 6 et 7 montrent pour quelques positions fixées de $\boldsymbol{k}$ la variation thermique de $I(\boldsymbol{k})$. Pour la région ferromagnétique ($T < T_c$).

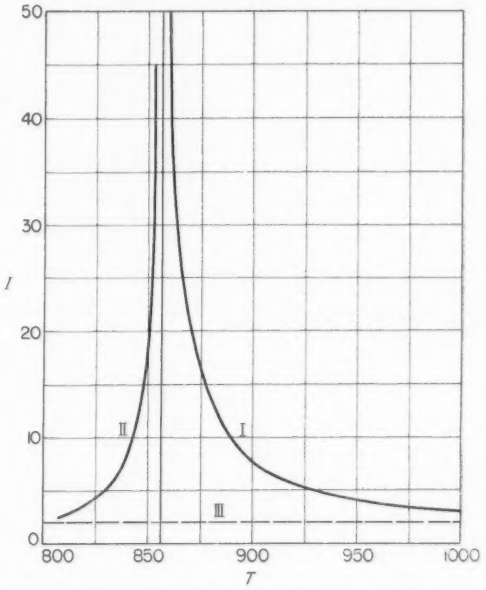


FIG. 5. Variations, en fonction de la température T, de l'intensité sur le nooud (111).

I : Variations de I pour $T > T_c$.
II : Variations de I_z pour $T < T_c$.
III : Asymptote pour $T \to \infty$.

les calculs sont faits par les mêmes techniques que précédemment, et nous n'en donnerons pas le détail ici ; les résultats sont également représentés sur les Figs. 5 et 7. Nous citerons enfin les valeurs du rapport $I_{\text{él}}/I_{\text{tot}}$ calculées d'après (51) dans la limite des hautes températures.

Tableau 1

hkl	000	200	220	222	400	111	333
$I_{\text{él}}/I_{\text{tot}}$	1	0	0,139	0,394	0,065	0,002	0,333

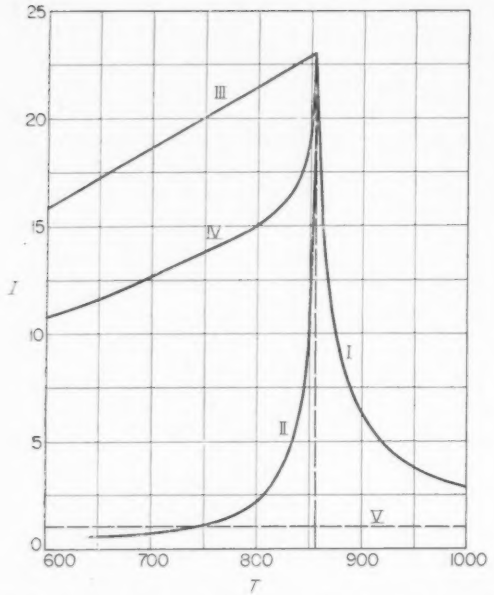


FIG. 6. Variations avec T de l'intensité I. On se place sur le cercle AIB, avec $\theta - \theta_B = 4^0$.

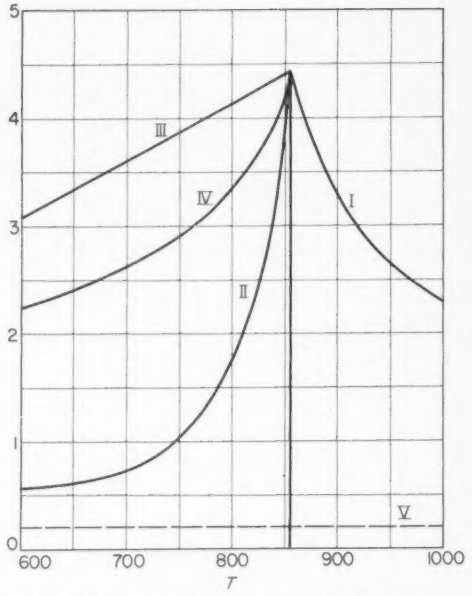


FIG. 7. Mêmes courbes que pour la Fig. 6 avec $\theta - \theta_B = 10^0$.

1 : Variations de I pour $T > T_c$
II : Variations de I_z pour $T < T_c$
III : Variations de I_x pour $T < T_c$
IV : Variations de $I = \frac{2}{3} I_x + \frac{1}{3} I_y$
V : Asymptote pour $T \to \infty$.

8. CONCLUSIONS GÉNÉRALES

Les différences de comportement, que nous avons mises en évidence, entre les substances magnétiques complexes et le cas simple ferromagnétique par VAN HOVE ont deux origines bien distinctes.

(1) Elles proviennent pour une part d'effets purement géométriques dus à la forme des déphasages $\boldsymbol{k}(\boldsymbol{r}_i - \boldsymbol{r}_j)$ entre atomes d'une même maille.

(2) Nous avons trouvé d'autre part, à propos des effets de freinage thermodynamique, des différences profondes de comportement statistique entre les antiferromagnetiques et les ferromagnétiques par exemple. Ces différences sont beaucoup moins essentielles quand on s'intéresse seulement aux propriétés statiques du domaine paramagnétique : par exemple la susceptibilité d'un antiferromagnétique pour un champ dont le signe alterne d'un sous réseau à l'autre diverge au point de transition T_c comme celle d'un ferromagnétique pour un champ homogène.[15] L'asymétrie de comportement dynamique provient de ce que la somme des aimantations est dans les deux cas une constante du mouvement, alors que les fluctuations importantes sont respectivement celles de la différence ou de la somme des aimantations.

Remarquons enfin que les résultats généraux que nous avons mis en évidence dans les sections 3–6 sont indépendants de l'approximation de champ moléculaire. Celle-ci limite par contre la précision des applications numériques, et pourrait être éventuellement améliorée par l'emploi de méthodes statistiques plus raffinées.[13,14]

Remerciements—Nous remercions M. A. HERPIN pour une discussion relative à ces questions, et Madame D. PILLARD qui a effectué les calculs numériques.

BIBLIOGRAPHIE

1. PALEVSKY H. et HUGHES D. J., *Phys. Rev.* **92**, 202 (1953).
2. SQUIRES G. L., *Proc. Phys. Soc. Lond.* **A67**, 248 (1954).
3. LOWDE R. D., *Proc. Roy. Soc.* **A235**, 305 (1956).

4. Wilkinson M. K. and Shull C. G., *Phys. Rev.* **103**, 516 (1956).
5. Galula M., Jacrot B. et Mangin J. P., *Colloque P. Weiss Strasbourg*. Edition du C.N.R.S. (1957). Galula M., Thèse, Paris (1958).
6. Lowde R. D., *Rev. Mod. Phys.* **30**, 69 (1958).
7. Macreynolds A. W. et Riste T., *Bull. Amer. Phys. Soc.* **1**, 133 (1956).
8. Riste T., *Congrès de Magnétisme, Grenoble* (1958).
9. MacReynolds A. W. et Riste T., *Phys. Rev.* **95**, 1161 (1954).
10. Riste T., Blinowski K. et Janik J., *J. Phys. Chem. Solids* **9**, 153 (1959).
11. Van Hove L., *Phys. Rev.* **95**, 1374 (1954).
12. de Gennes P. G. et Herpin A., *C.R. Acad. Sci. Paris*, **243**, 1611 (1956).
13. Elliott R. J. et Marshall W., *Rev. Mod. Phys.* **30**, 75 (1958).
14. de Gennes P. G., *Colloque P. Weiss Strasbourg*. Editions du C.N.R.S. (1957) et These, Université de Paris (1957).
15. de Gennes P. G., *J. Phys. Chem. Solids* **6**, 43 (1958).
16. Yvon J., *Rev. Sci. Paris* **11**, 662 (1939).
17. de Gennes P. G. et Friedel J., *J. Phys. Chem. Solids* **4**, 71 (1958).
18. Dyson F. J., *Phys. Rev.* **102**, 127 (1956).
19. de Gennes P. G., *J. Phys. Chem. Solids* **4**, 223 (1958).
20. Cribier D., Galula M., Jacrot B. et Sarma G., *C.R. Acad. Sci. Paris* **248**, 1631 (1959).
21. Néel L., *Ann. Phys. Paris* **3**, 137 (1948).

VOL.
13
1960

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 13. pp. 28–32. Printed in Great Britain.

CRYSTAL CHEMICAL AND MAGNETIC STUDIES OF GARNET SYSTEMS

$\{YCa_2\}[M_2^{4+}](Fe_3)O_{12}$—$\{Y_3\}[Fe_2](Fe_3)O_{12}$, M = Zr or Hf

S. GELLER, R. M. BOZORTH, C. E. MILLER and D. D. DAVIS

Bell Telephone Laboratories, Inc. Murray Hill, New Jersey

(*Received* 12 *August* 1959)

Abstract—A complete solid-solution range exists in the system $\{YCa_2\}[Zr_2](Fe_3)O_{12}$—$\{Y_3\}[Fe_2](Fe_3)O_{12}$. The Zr^{4+} ion is even larger than the Sn^{4+} ion and in this system apparently prefers only the octahedral sites. The behavior of the plot of $n_B(\infty,0)$ vs. composition is similar to that of the $Ca_3Fe_2Sn_3O_{12}$—$Y_3Fe_2Fe_3O_{12}$ system and corroborates the interaction model for garnets recently derived by GILLEO. A small spontaneous magnetization is observed in the garnets $\{YCa_2\}[Zr_2](Fe_3)O_{12}$ and $\{YCa_2\}[Zr_2](Me_{0.5}Fe_{2.5})O_{12}$, Me = Al and Ga, indicating the establishment of magnetic sublattices among the tetrahedral sites.

The garnet $\{YCa_2\}[Hf_2](Fe_3)O_{12}$ is readily formed. The Hf^{4+} ion, which is very nearly the same size as that of Zr^{4+}, also prefers the octahedral site.

1. INTRODUCTION

THE garnet structure belongs to space-group *Ia3d* with cations in 16*a*, 24*c* and 24*d* and oxygen (or fluorine) ions in 96*h*. If the distribution of the ions is known, the chemical formula of one unit is written $\{A_3\}[B_2](C_3)O_{12}$, where { } represents a *c* or dodecahedral site, [ ] an *a* or octahedral site and ( ) a *d* or tetrahedral site. Extensive crystal-chemical and magnetic studies on the garnets have been made in these Laboratories. Work on Group IVb tetravalent ions Si^{4+}, Ge^{4+} and Sn^{4+} has shown[1, 2] that these ions substitute for Fe^{3+} in yttrium–iron garnet, $\{Y_3\}[Fe_2](Fe_3)O_{12}$, in a fashion similar to that of the nonmagnetic trivalent ions.[3] The Si^{4+} ion shows a great preference for tetrahedral sites, although it does appear to enter octahedral sites to a small extent; Ge^{4+} ion also prefers the tetrahedral sites, but goes into octahedral sites to a much greater extent than does Si^{4+}; Sn^{4+} shows a very great preference for octahedral sites.

In work on new synthetic garnets,[4] it has been indicated that Ti^{4+} and Zr^{4+} go into the garnets and show a preference for octahedral sites. Because Hf^{4+} has very nearly the same size as Zr^{4+} and because hafnium chemistry is very closely analogous to that of zirconium, one might expect the two to act similarly in the garnets. This paper will be mainly concerned with the substitution of Zr^{4+} and Hf^{4+} for Fe^{3+} ion in yttrium–iron garnet. Work on Ti^{4+} substitution has not yet been completed.

2. EXPERIMENTAL

(a) *Preparation of specimens*

(i) $\{Y_{3-y}Ca_y\}[Zr_yFe_{2-y}](Fe_3)O_{12}$ system. Required amounts of oxides were thoroughly mixed, ground together in an agate mortar and compacted into pellets. These were fired at temperatures of 1350–1450°C in air for 2 hr. Most of the specimens were reground and refired three to six times until X-ray photographs of the specimens showed sharp lines and indicated absence of unwanted phases.

In several cases it was found that 16-hr firings at 1300°C gave very dense specimens, which in turn gave very sharp photographs. However, there was evidence, both crystallographic and magnetic, that these long firings were causing some decomposition; in particular, the Zr^{4+} ion was leaving the garnet phase.

(ii) $\{Y_{3-y}Ca_y\}[Hf_yFe_{2-y}](Fe_3)O_{12}$ system. In this system the end-member, $y = 2$, was readily made.

VOL. 13 1960

Firing conditions were: 1 hr at 1400°C followed by 17 hr at 1300°C. A specimen for $y = 1{\cdot}0$ was obtained, but the X-ray photographs of this composition were not sharp.

(b) *Diffraction data*

X-ray powder photographs were taken of all specimens with Straumanis-type Norelco cameras of 114·6 mm dia.; CrK radiation was used.

(c) *Magnetic measurements*

Magnetic measurements were made as described elsewhere[1] in fields up to 12,500 Oe, and at temperatures of 1·3–300°K. In the cases of small zirconium substitution, saturation magnetization was determined by extrapolation to $T = 0$ and $H = \infty$, using the $1/H^2$ relation from $H = 8000$ to 12,000 or 12,500 Oe. For $y \geqslant 1{\cdot}4$, the proper method of extrapolation was not entirely clear, and in these cases the extrapolation was performed both as above and to $H = 0$ (see Table 2).

3. CRYSTALLOGRAPHIC DATA

The lattice constants of the various specimens, $\{Y_{3-y}Ca_y\}[Zr_yFe_{2-y}](Fe_3)O_{12}$, are listed in Table 1 and plotted in Fig. 1.

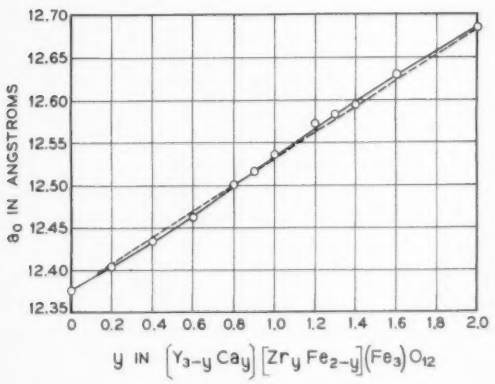


FIG. 1. Lattice constant vs. composition of garnet system $\{Y_{3-y}Ca_y\}[Zr_yFe_{2-y}](Fe_3)O_{12}$.

It seemed that the more heavily Zr^{4+}-substituted yttrium–iron garnets form more readily than those with lesser Zr^{4+} substitution. In the case of the Hf^{4+} ion substitution the difficulty was magnified. A sharp powder photograph of the end-member $\{YCa_2\}[Hf_2](Fe_3)O_{12}$ gave a lattice constant of 12·670±0·004 Å for this specimen, to be compared with 12·684±0·004 Å (Table 1) for $\{YCa_2\}[Zr_2](Fe_3)O_{12}$.

Table 1. *Lattice constants of the system* $\{Y_{3-y}Ca_y\}[Zr_yFe_{2-y}](Fe_3)O_{12}$

y	Lattice constant (Å)
0	12·376±0·004
0·2	12·404
0·4	12·434
0·6	12·463
0·8	12·501
0·9	12·516
1·0	12·535
1·2	12·572
1·3	12·583
1·4	12·594
1·6	12·630
2·0	12·684

Because it appeared that all of the Zr^{4+} goes into octahedral sites, it was of interest to introduce a controlled amount of nonmagnetic ion into the tetrahedral sites of the garnet $\{YCa_2\}[Zr_2](Fe_3)O_{12}$. The following two were made:

$\{YCa_2\}[Zr_2](Fe_{2.5}Al_{0.5})O_{12}$ $a_0 = 12{\cdot}650 \pm 0{\cdot}004$ Å

$\{YCa_2\}[Zr_2](Fe_{2.5}Ga_{0.5})O_{12}$ $a_0 = 12{\cdot}680 \pm 0{\cdot}004$ Å

The crystallographic evidence for the complete preference of Zr^{4+} and Hf^{4+} for the octahedral sites is as follows. First, these ions are larger than the Sn^{4+} ion. Quantitative X-ray diffraction measurements have shown that the specimen with nominal formula $YCa_2Sn_2Fe_3O_{12}$ has the ionic distribution $\{YCa_2\}[Sn_{1.8}Fe_{0.2}](Sn_{0.2}Fe_{2.8})O_{12}$. Secondly, qualitative examination of the powder intensities corroborates this deduction: One of the most sensitive reflections is the 444, which decreases in intensity with substitution of a higher-atomic-number element in the octahedral site. As in the case of $\{YCa_2\}[Sn_{1.8}Fe_{0.2}](Sn_{0.2}Fe_{2.8})O_{12}$, this reflection is very weak in the powder photograph of $\{YCa_2\}[Zr_2](Fe_3)O_{12}$ and unobservable in the powder photograph of

$\{YCa_2\}[Hf_2](Fe_3)O_{12}$.

The latter observation is also evidence that the

Zr^{4+} or Hf^{4+} ions are not entering the dodecahedral sites to any appreciable extent.

4. MAGNETIC DATA

Typical curves of magnetization vs. temperature are shown in Figs. 2 and 3. The magnetization curve for low zirconium content comes into the ordinate with zero slope (Fig. 2) and that for

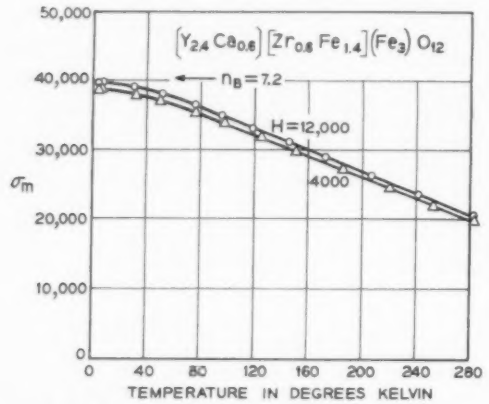


FIG. 2. Magnetization vs. temperature of garnet $\{Y_{2\cdot4}Ca_{0\cdot6}\}[Zr_{0\cdot6}Fe_{1\cdot4}](Fe_3)O_{12}$.

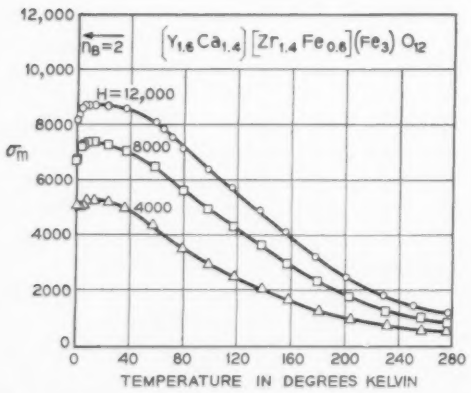


FIG. 3. Variation of magnetization with temperature of $\{Y_{1\cdot6}Ca_{1\cdot4}\}[Zr_{1\cdot4}Fe_{4\cdot6}](Fe_3)O_{12}$.

higher zirconium content appears to reach a maximum at about 10°K and then decreases, thus cutting the ordinate with a positive slope (Fig. 3). This behaviour is very similar to that of the tin-substituted yttrium–iron garnet system.[1, 2]

5. DISCUSSION

The results of this work indicate that the Group IVb ions Zr^{4+} and Hf^{4+} go entirely into octahedral sites when substituted into the garnets. This is in accord with the rules established for substitution of ions into the garnet structure, namely: (1) the octahedral and tetrahedral sites prefer spherical (or pseudospherical) ions; (2) of these two sites, the octahedral always allows more space to an occupying ion, the larger ion preferring the larger site. The first rule is the stronger one; thus Cr^{3+} which has a pseudospherical ground state only in an octahedral crystal field, substitutes for octahedral Fe^{3+} ion in $\{Y_3\}[Fe_2](Fe_3)O_{12}$[3] even though the Cr^{3+} ion is smaller[5] than the Fe^{3+} ion. However, the maximum amount of Cr^{3+} ion that can be

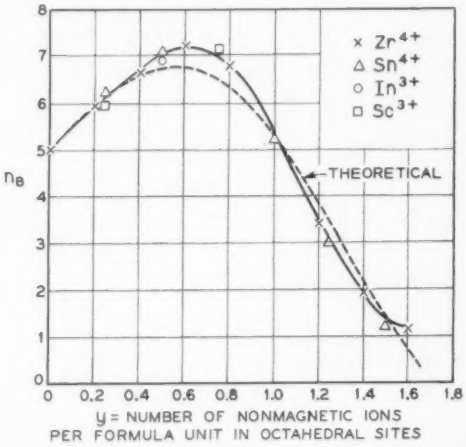


FIG. 4. Theoretical and experimental variation of magnetic moment (at 0°K) with composition, in $\{Y_{3-y}Ca_y\}[Zr_yFe_{2-y}](Fe_3)O_{12}$.

substituted for Fe^{3+} ion in $\{Y_3\}[Fe_2](Fe_3)O_{12}$ is small ($y = 0{\cdot}25$);[3] larger amounts would apparently be incompatible with the structural requirements. It is of interest to note, however, that the garnet $Ca_3Cr_2Si_3O_{12}$ does exist.[6] In this case, the ratio of Cr—O to Si—O distances is apparently in accord with the structural requirements. Thus the rules have structural limitations imposed on them.

The Zr^{4+} and Hf^{4+} ions are of about equal size and have spherical electronic configurations: $4s^24p^6$ and $5s^25p^6$ respectively. These ions are larger than the Sn^{4+} ion, which, as has been shown, greatly prefers octahedral coördination in the

garnets. However, our results indicate that, unlike the Sn^{4+} ion which will go into the tetrahedral sites, the Zr^{4+} and Hf^{4+} ions do not.

Not only did the semiquantitative analysis of the X-ray diffraction powder intensities indicate the preference of the Zr^{4+} and Hf^{4+} ions for the octahedral sites, but the curve (Fig. 4) obtained from measurement of the 0°K (extrapolated) saturation magnetization vs. composition practically superposes on that of the tin-substituted yttrium–iron garnet system.[2]

As a result of the recent work on the latter system, GILLEO has worked out a model for superexchange interaction in garnets containing randomly incomplete linkages,[7] which result from the substitution of non-magnetic for magnetic ions. If we justifiably assume that all of the Zr^{4+} ions occupy octahedral sites in the various specimens, the saturation magnetization (excluding very low-temperature effects, see Refs. 2 and 7 and later discussion) at 0°K of a garnet with formula $\{Y_{3-y}Ca_y\}[Zr_yFe_{2-y}](Fe_3)O_{12}$ is given by

$$n_B = 5\{3[1-4(y/2)^3+3(y/2)^4]-2(1-y/2)\}\mu_B$$

Values of n_B have been calculated for the specimens measured and are compared with the observed values in Table 2 and Fig. 4.

As a consequence of GILLEO's recent work,[7] the Curie temperatures may also be predicted and are given by

$$T_C(y/2) = \frac{5(1-y/2)[1-4(y/2)^3+3(y/2)^4]}{2(1-y/2)+3[1-4(y/2)^3+3(y/2)^4]} \times 545.$$

The Curie temperatures are not easily accurately determined from the data taken and thus only two observed Curie temperatures are listed in Table 2. However in Fig. 5 the data thus far obtained for specimens in which the nonmagnetic ions Zr^{4+}, Sn^{4+}, Sc^{3+}* and In^{3+}* occupy only octahedral sites, are plotted and compared with the curve

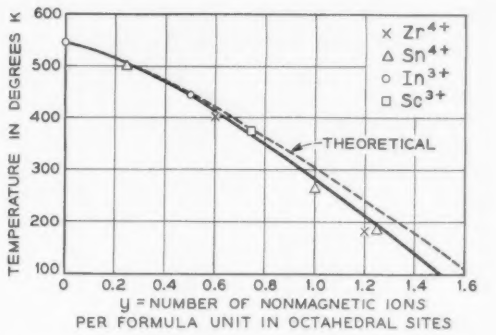


FIG. 5. Theoretical and experimental variation of Curie temperatures with composition.

obtained from the formula (see Tables 2 and 3 for actual values plotted).

Table 2. Saturation magnetizations at 0°K and some Curie temperatures in the system $\{Y_{3-y}Ca_y\}[Zr_yFe_{2-y}](Fe_3)O_{12}$

y	$\lvert n_B\rvert$(obs.) (μ_B)		n_B(calc.) (μ_B)	T_C(obs.) (°K)	T_C(calc.) (°K)
0·0	5·0		5·00	545	545
0·2	5·9		5·94		510
0·4	6·7		6·59		469
0·6	7·2		6·75	400	421
0·8	6·8		6·31		366
1·0	5·3		5·31		306
1·2	3·4		3·87	180	241
1·4	2·0	0·6*	2·22		173
1·6	1·2	0·2*	0·71		105
2·0	0·7	0·2*	0·00		

* Extrapolated to $T = 0$ and $H = 0$. Note that others are extrapolated to $T = 0$ and $H = \infty$ (see text).

Both the observed n_B vs. composition and T_C vs. composition curves show some deviation from the curves given by GILLEO's interaction model. For the higher values of y the measured values of both n_B and T_C are in some doubt; therefore it is not known just how significant the deviation is. However, for the lower values of y, i.e. $y \leqslant 0{\cdot}8$, there is a significant deviation, although the largest, for $y = 0{\cdot}8$, is only 8 per cent. There is little doubt that the theoretical model is an excellent one, however.

* In Ref. 3, the results indicate various percentages of In^{3+} and Sc^{3+} entering tetrahedral sites when these ions are substituted for Fe^{3+} ion in yttrium-iron garnet. However, these results were obtained prior to the work on the tin-substituted yttrium–iron garnet which led to a better understanding of the results of substitution of other ions for the Fe^{3+} ion. It is now concluded that in the specimens given in Ref. 3, all the In^{3+} and Sc^{3+} ions occupy octahedral sites.

VOL. 13 1960

Table 3. Saturation magnetizations at 0°K and some Curie temperatures in the systems $\{Y_{3-y}Ca_y\}[Sn_yFe_{2-y}](Fe_3)O_{12}$,[2] $\{Y_3\}[In_yFe_{2-y}](Fe_3)O_{12}$[3, 7] and $\{Y_3\}[Sc_yFe_{2-y}](Fe_3)O_{12}$[3, 7]

Ion	y	$\lvert n_B\rvert$(obs.) (μ_B)	n_B(calc.) (μ_B)	T_C(obs.) (°K)	T_C(calc. (°K)
Sn^{4+}	0·25	6·25	6·15	498	503
	0·5	7·1	6·75		446
	1·0	5·2	5·30	260	306
	1·25	2·9	3·50	180	224
	1·5	1·2	1·42		138
In^{3+}	0·5	6·9	6·75	444	446
Sc^{3+}	0·25	5·94	6·15	500	503
	0·75	7·2	6·50	375	381

In two previous papers,[1, 2] it was noted that tetrahedral–tetrahedral (d–d) magnetic interactions in the garnets were directly observed. In order to corroborate these observations further, small amounts of Al^{3+} and Ga^{3+} ions were substituted for iron in the garnet $\{YCa_2\}[Zr_2](Fe_3)O_{12}$. These ions almost surely go into tetrahedral sites and may be written: $\{YCa_2\}[Zr_2](Me_zFe_{3-z})O_{12}$, where Me = Al or Ga. We made the garnets $z = 0{\cdot}5$ (see Section 3); both of these showed a slight parasitic ferrimagnetization of 0·1 μ_B at 0°K (extrapolated to $H = 0$). This indicates that there are at least two magnetic sublattices among the tetrahedral sites in the garnet $\{YCa_2\}[Zr_2](Me_zFe_{3-z})O_{12}$ and that the magnetic structure is one not permitting strict antiparallel alignment of the moments of the ions in the two sublattices.

In the system containing Sn^{4+} it was observed that for high Sn^{4+} contents, the n_B vs. T curves at larger fields appeared to have maxima at very low temperatures.[1, 2] This is also true of the system $\{Y_{3-y}Ca_y\}[Zr_yFe_{2-y}](Fe_3)O_{12}$ for large y; (see, for example, Fig. 3). In all cases the $n_B(T = 0°K)$ values reported (Fig. 4) were obtained by extrapolating the lower n_B values (that is, where the curves actually intercepted the ordinate) to $H = \infty$ or in some cases $H = 0$ as indicated. The higher values of n_B obtained by extrapolating from above the maxima do not in any case exceed the reported values by more than 0·4μ_B.

It may well be that the higher values should be the ones to be compared with GILLEO's model, because his model does not include very low-temperature effects: It has already been mentioned in an earlier paper[2] that the behavior of these curves is indicative of a low-temperature magnetic transition. Such a transition would almost surely be of higher than first-order, the magnetic structure changing gradually from one to another over a temperature range. Such transitions have been observed in intermetallic compounds; there appears to be such a transition in the compound Co_5Gd;[8, 9] in Ni_3Mn, there appears to be a ferrimagnetic–ferromagnetic transition at very low temperatures.[10]

Acknowledgements—The authors wish to thank A. J. WILLIAMS for aid with the magnetic measurements.

REFERENCES

1. BOZORTH R. M. and GELLER S., *J. Phys. Chem. Solids* **11**, 263 (1959).
   See also BOZORTH R. M., GELLER S. and MILLER C. E., *Bull. Amer. Phys. Soc.* **4**, 241 (1959).
2. GELLER S., BOZORTH R. M., GILLEO M. A. and MILLER C. E., *J. Phys. Chem. Solids* **12**, 111 (1960).
   (See also *Bull. Amer. Phys. Soc.* **4**, 241 (1959).
3. GILLEO M. A. and GELLER S., *Phys. Rev.* **110**, 73 (1958).
4. GELLER S. and MILLER C. E., *Acta Cryst., Camb.* In press. See also *Amer. Cryst. Assoc. Abstr.* No. E–1 (July 1959).
5. GELLER S., *Acta Cryst., Camb.* **10**, 248 (1957).
6. GELLER S. and MILLER C. E., *Amer. Min.* **44**, 445 (1959) and references in this note.
7. GILLEO M. A., *J. Phys. Chem. Solids* **13**, 33 (1959).
8. NESBITT E. A., WERNICK J. H. and CORENZWIT E., *J. Appl. Phys.* **30**, 365 (1959).
9. WERNICK J. H. and GELLER S., *Acta Cryst., Camb.* **12**, 1662 (1959).
10. KOUVEL J. S., GRAHAM C. D. and BECKER J. J., *J. Appl. Phys.* **29**, 518 (1958).

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 13. pp. 33–39. Printed in Great Britain.

SUPEREXCHANGE INTERACTION IN FERRIMAGNETIC GARNETS AND SPINELS WHICH CONTAIN RANDOMLY INCOMPLETE LINKAGES

M. A. GILLEO

Lockheed Aircraft Corporation, Missiles and Space Division, Sunnyvale, California

(*Received* 10 *August* 1959)

VOL. 13 1960

Abstract—The variation of saturation moment, Curie temperature and susceptibility of ferrimagnetic compounds as a result of substitution of nonmagnetic ions for magnetic ions, can be understood to good approximation on the basis of the random distribution of incomplete superexchange interactions. The two premises required for the understanding are that: (1) a magnetic ion actively participates in ferrimagnetism only if it interacts with two or more magnetic ions in different coördination and (2) the Curie temperature is determined by the number of complete interactions per active magnetic ion per formula unit. The requirement that a magnetic ion interact with at least two other magnetic ions in different coördination is implicit in the fact that ferrimagnetism is a coöperative phenomenon. Recent X-ray diffraction and magnetic studies of the yttrium–iron–calcium–iron–tin garnet system have provided the first firm basis for this understanding.

1. INTRODUCTION

THE importance of superexchange interaction of magnetic ions to the understanding of the magnetic properties of oxides of transition metals was first recognized by KRAMERS[1]. Later NÉEL[2] demonstrated that ferrimagnetism results from negative superexchange (antiferromagnetic) interaction. In all cases the unit of interaction is a linkage of the M_a–O–M_b type. Excited states of the oxygen ion which separates the two transition-metal ions couple their spins through interaction with their *d*-electrons. The theory of superexchange interactions has been brought to its most advanced state at the present time by the work of ANDERSON[3].

The Curie temperatures of oxides of iron of ideal crystal-chemical composition* have been found to be closely related to the number n of Fe^{3+}–O^{2-}–Fe^{3+} linkages per Fe^{3+} ion per formula unit.[4] Therefore, an energy of interaction of kT_C/n may be assigned to each linkage. The chains of such linkages are indefinitely long in a compound of ideal crystal-chemical composition. However, for compounds in which part of the iron ions has been replaced by nonmagnetic ions, the crystal-chemical composition is no longer ideal, so that the number of *active* linkages will vary statistically from one magnetic ion to the other.

In the case of substituted yttrium–iron garnets[5] an earlier attempt was made to interpret the observed variation of magnetic moment and Curie temperature on the basis of the change in the number of active exchange linkages per magnetic ion. The assumption made was that a magnetic ion which is linked only with nonmagnetic ions through the surrounding oxygen ions can not participate, as other magnetic ions do, in the ferrimagnetism of the compound. In this respect interaction between magnetic ions in crystallographically equivalent sites was not considered, for it is known that the geometry of such linkages is unfavourable to superexchange interaction in the garnet structure.[4, 6]

The data obtained from the recently synthesized solid solutions of yttrium–iron garnet (YIG) with $YCa_2Sn_2Fe_3O_{12}$(CIS) and $YCa_2Zr_2Fe_3O_{12}$ (CIZ) garnets[7–9] now make it possible to extend greatly

* A compound of ideal crystal-chemical composition is here considered to be one in which a given crystallographic position is occupied, without exception, by an ion of a single kind and valence.

our understanding of the dependence of magnetic moment and Curie temperature upon the statistical distribution of active linkages among the magnetic ions which participate in the ferrimagnetism of the compound. It is clear that the original approximation[5] in which account was taken only of magnetic ions excluded by lack of interaction with any other magnetic ions of another coördination, is no more than a first-order approximation; to account adequately for the data now available, it should be carried to a higher order.

2. INTERACTION STATISTICS

(a) *Magnetic Moments*

For substituted yttrium–iron garnets it is convenient to use the following formula unit:

$$\{Y_{3-y-z}Me_{y+z}\}[Fe_{2-y}M_y](Fe_{3-z}M_z)O_{12}.$$

In a unit cell of space-group O_h^{10}—$Ia3d$, to which the yttrium–iron garnet structure belongs,[6] the ions of the formula unit which are in dodecahedral sites, 24(c), are enclosed by { }, those in octahedral sites, 16(a), by [ ], and those in tetrahedral sites, 24(*d*), by ( ). If M is a quadrivalent ion, such as Si^{4+}, Sn^{4+} or Zr^{4+}, Me is a divalent ion, such as Ca^{2+};[7–9] if M is trivalent, so also is Me, which then is Y^{3+} or a rare-earth ion.[5, 10] Very many other possibilities are available.[11–14]

It is known from structure studies[6] that an octahedral and a tetrahedral ion are linked with each other only once through an oxygen ion, that is to say that each oxygen ion is coördinated with only one each of the octahedral and tetrahedral ions. Therefore, for each of two octahedral iron ions there are six linkages with tetrahedral ions for a total of twelve; and for each of three tetrahedral ions there are four linkages with octahedral ions for a total of twelve. Consequently, there are 24/5 linkages per iron ion in yttrium–iron garnet. The magnetic moment is equal to that of the three tetrahedral Fe^{3+} ions of 5 μ_B each minus that of two octahedral Fe^{3+} ions, which leaves 5 μ_B as the net moment per O_{12} formula unit.[5]

However, when iron ions are replaced by nonmagnetic ions M in amounts y and z, a number of linkages are incomplete. Magnetic ions which interact with no other magnetic ions of another coördination must be excluded from participation in ferrimagnetism.[5] In addition, with a large replacement of iron ions by nonmagnetic ions, the possibility that a magnetic ion may interact with only one other magnetic ion in another coördination must be taken into account. An ion with a single linkage cannot be active in a coöperative interaction with the remainder of the ions, as is required in ferrimagnetism, because it cannot relay the orientation information contained in the single linkage to any other magnetic ions; that is to say that the interaction chain through a magnetic ion must be uninterrupted in order that the ion may actively participate in ferrimagnetism.

Consequently, a single linkage should be discarded and the ion thus linked also excluded from participation in ferrimagnetism. NÉEL[15] first observed that the dependence of the saturation magnetization of zinc-substituted nickel ferrite upon zinc content could be qualitatively explained by the assumption that an octahedral magnetic ion with less than two linkages with tetrahedral magnetic ions did not contribute to the ferrimagnetism.

Let k_o and k_t be the fractions of octahedral and tetrahedral Fe^{3+} ions which are replaced by nonmagnetic ions, i.e.

$$k_o = y/2,$$
$$k_t = z/3.$$

The probability $p_n(m)$ that an ion of coördination n be linked with m ions, each of probability $(1-k)$ is

$$p_n(m) = \binom{n}{m}k^{n-m}(1-k)^m,$$

where $\binom{n}{m}$ is the binomial coefficient,

$$n!/m!(n-m)!$$

Therefore, the probability E that an ion be linked with none, or at most one, of the magnetic ions is

$$E = \sum_{m=0}^{1} p_n(m) = nk^{n-1}-(n-1)k^n.$$

If the fraction E of the magnetic ions in a specified site is excluded from participation in the ferrimagnetic structure of the compound, then only the portion $(1-E)(1-k)$ of the ions in that site can contribute to the ferrimagnetism. For a compound of garnet structure the moments $n_B^{(o)}$, $n_B^{(t)}$ of the octahedral and tetrahedral sublattices are

$$n_B^{(o)} = 2\mathcal{M}_o(1-k_o)[1-E_o(k_t)],$$
$$n_B^{(t)} = 3\mathcal{M}_t(1-k_t)[1-E_t(k_o)],$$

where

$$E_o(k_t) = 6k_i^5 - 5k_i^6,$$

$$E_t(k_o) = 4k_o^3 - 3k_o^4,$$

and $\mathcal{M}_o, \mathcal{M}_t$ are the moments in Bohr magnetons, μ_B, of the magnetic ions.

For an iron garnet with Fe^{3+} ions of 5 μ_B moment, the moment per O_{12} formula unit is

$$n_B(k_o,k_t) = n_B^{(t)} - n_B^{(o)}$$
$$= 3{\cdot}5(1-k_t)[1-E_t(k_o)] - 2{\cdot}5(1-k_o)[1-E_o(k_t)]\mu_B.$$

The moment derived in this way is not the true spontaneous magnetization which would be observed with zero field at absolute zero of temperature; it corresponds to the apparent moment which would be deduced from measurements made in weak fields down to, say, 10°K. Weaker interactions which are important at absolute zero in zero field have not been considered and will be discussed later (see also Ref. 7).

(b) *Curie temperature*

For magnetic oxides of iron of ideal crystal-chemical composition, it has been found that the Curie temperature corresponds to the number of interactions per magnetic ion per formula unit.[4] Therefore, reduction of the number of interactions by substitution of nonmagnetic ions for Fe^{3+} ions should lower the Curie temperature.

The number of primary interest to us for the determination of the Curie temperature is, therefore, the number of active interactions per magnetic ion. Because each linkage involves two magnetic ions, we would then count just twice as many interactions as linkages. For twelve oxygen ions per formula unit, each of which is coördinated with one octahedral and one tetrahedral magnetic ion,[6] we would have a maximum of twenty-four interactions. However, with substitution of a fraction k_i of non magnetic ions for iron ions of kind i, a fraction $(1-k_i)$ of the ions is magnetic and of these a fraction $[1-E_i(k_j)]$ is active.

Accordingly, the number of active interactions (most easily counted from the oxygen position as a product of the probabilities of each linkage half) is

$$n(k_o,k_t) = 24(1-k_o)[1-E_o(k_t)](1-k_t)\times$$
$$\times[1-E_t(k_o)]$$

for an O_{12} formula unit.

The number N of magnetic ions which actively participate in the ferrimagnetism is

$$N(k_o,k_t) = 2(1-k_o)[1-E_o(k_t)] +$$
$$+3(1-k_t)[1-E_t(k_o)].$$

The Curie temperature $T_C(k_o,k_t)$ for a substituted garnet is related to that of yttrium–iron garnet, $T_C(0,0) = 545°K$,[5] by the relation

$$T_C(k_o,k_t) = \frac{5}{24}\frac{n(k_o,k_t)}{N(k_o,k_t)}T_C(0,0),$$

where 24/5 is the number of interactions per magnetic ion per formula unit.

3. COMPARISON WITH EXPERIMENT

The comparison of calculated and observed values of n_B and T_C computed for substituted yttrium–iron garnets on the basis of ion distribution, which is known from X-ray or neutron-diffraction data, shows quite good agreement, particularly for octahedral substitutions (Table 1). Where the distribution is not known from independent observation, it has been based on the best available information.

For the most part results previously reported[5] for substitutions which take place principally in tetrahedral sites are only slightly affected in interpretation by the higher order of approximation afforded by the more elaborate analysis given here; the substitution fraction k_t is too small to bring about appreciable departure from the earlier first order approximation.

When substitution takes place primarily in octahedral sites, the effects on the magnetic moment and Curie temperature of yttrium–iron garnet are much more pronounced than for tetrahedral substitution: first, the moment passes through a maximum before declining to zero, and second, the effects of substitution are more rapid because of the smaller number of octahedral ions present in a formula unit.

The substitution fraction at which the maximum value of $n_B(k_o,0)$ occurs corresponds to the point at which $\partial n_B(k_o,0)/\partial k_o = 0$ which is $k_o = 0{\cdot}277$ (i.e. $y = 0{\cdot}554$). The maximum moment which can be achieved by non magnetic ion substitution for a given fraction k of the iron, where $5k = 2k_o + 3k_t$, occurs for total octahedral substitution ($k_o = (5/2)k$, $k_t = 0$) for $0 \leqslant k \leqslant 0{\cdot}165$; at $k = 0{\cdot}165$, $\partial n_B(k_o,0)/\partial k_o|_k = 0$; for $k > 0{\cdot}165$,

Table 1. Observed and calculated magnetic moments and Curie temperatures for substituted yttrium–iron garnet

Formula unit	k_o	k_t	n_B (obs.) (μ_B)	n_B (calc.)(μ_B)	T_C (obs.)(°K)	T_C (calc.)(°K)	Ref.
$Y_3Fe_5O_{12}$	0	0	5	5	545 ± 3	545	5
$Y_{11/4}Ca_{1/4}Fe_{19/4}Sn_{1/4}O_{12}$	0·125*	0	6·25	6·14	498 ± 3	503	7
$Y_3Fe_{19/4}Sc_{1/4}O_{12}$	0·125†	0	5·94	6·14	500 ± 3	503	5
$Y_{5/2}Ca_{1/2}Fe_{9/2}Sn_{1/2}O_{12}$	0·250*	0	7·1	6·74	—	446	7
$Y_3Fe_{9/2}In_{1/2}O_{12}$	0·250†	0	6·9	6·74	444 ± 3	446	5
$Y_3Fe_{17/4}Sc_{3/4}O_{12}$	0·375†	0	7·2	6·48	375 ± 3‡	381	5
$Y_2CaFe_4SnO_{12}$	0·500*	0	5·2	5·31	260 ± 20	306	7,8
$Y_{7/4}Ca_{5/4}Fe_{15/4}Sn_{5/4}O_{12}$	0·625*	0	2·9	3·47	180 ± 20	224	7
$Y_{3/2}Ca_{3/2}Fe_{7/2}Sn_{3/2}O_{12}$	0·750*	0	1·2	1·44	—	138	7
$YCa_2Fe_3Sn_2O_{12}$	0·9*	0·067	0·25	−0·27	—	41	7

* These values have been quantitatively established by X-ray diffraction intensity data in Ref. 7.

† These values are based on data of Ref. 5. Values given there were derived from magnetic data on the basis of the simple model employed. The X-ray diffraction studies (Ref. 7) and ionic-size considerations (Ref. 7) and the improved model derived in this paper are the basis for revision.

‡ The value of $T_C = 365$°K given in Table II of Ref. 5 is in error and should read 375°K; the curve from which it is derived (Fig. 4) is correct.

$\partial n_B(k_o,0)/\partial k_o|_k<0$, so that the largest possible moment is realized for $k_t>0$.

Although the agreement between the observed[5,7,8] and calculated moments is quite good, adequate account has not been taken of the contribution of magnetic ions which have been considered not to participate in ferrimagnetism. The moments of these ions must become ordered at very low temperature. For a small fraction $E_t(k_o)$ of tetrahedral ions excluded from interaction, it is probable that at absolute zero temperature the fraction $(1-k_t)k_0^4$ of the tetrahedral ions which is magnetic and which interacts with no octahedral ions is oriented antiparallel with the ferrimagnetic tetrahedral ions as a consequence of weak interaction[16] with them. Because this interaction is weak, the influence of an external field H, such that $\mu H > kT$ for the excluded ions, would cause the moment of these tetrahedral ions to add to the ferrimagnetic moment, whereas in zero field at very low temperature it would subtract. The case for those tetrahedral ions which have a single linkage with an octahedral ion is less determinate. The weak interaction[16] of these tetrahedral ions with other tetrahedral ions opposes the single ferrimagnetic interaction so that the orientation of such ions in zero field is difficult to predict. Accordingly, the temperatures and magnetic fields used in the determination of the spontaneous magnetization by extrapolation to zero temperature could seriously affect the result.

Data reported by GELLER *et al.*[7] for $Ca_{1.5}Y_{1.5}Fe_{3.5}Sn_{1.5}O_{12}$ show that the magnetic moment for temperatures as low as 1·3°K is not approaching a final value at 0°K with zero slope, as it ultimately must. For Fe^{3+}, with a moment of 5 μ_B in a field of 12,000 Oe, $\mu H \simeq kT$ at 4°K. Therefore, the positive slope of σ_m vs. T observed at 12,000 Oe below 10°K implies that the Néel temperature for interaction between tetrahedral iron ions is >4°K.

Under these circumstances the problem of spontaneous magnetization becomes much more complicated because of the increased number of interactions which must be taken into account. In particular, these additional interactions will very probably lead to conditions under which the moments may no longer be simply antiparallel in magnetic lattices of ions in crystallographically different positions, and parallel for ions in the same positions.[7] The treatment of such a problem by YAFET and KITTEL[17] indicates the manner in which additional interactions may influence the magnetization at very low temperatures.

VOL. 13 1960

For a ferrimagnetic garnet with $k = 3/10 = (2/5)k_o$, the value of $E_t(k_o)$ is 0·74, so that $3 \times 0{\cdot}74 = 2{\cdot}2$ iron ions in tetrahedral sites are excluded from participation in ferrimagnetism except at very low temperatures (say, $<10°$K). Therefore, these excluded ions would be responsible for a substantial contribution to the susceptibility of such a compound for which the moment may be represented by $\sigma = \sigma_0 + \chi H$. Susceptibility of a magnitude appropriate to 2·2 free Fe^{3+} ions per formula unit is observed for $Ca_{1{\cdot}5}Y_{1{\cdot}5}Fe_{3{\cdot}5}Sn_{1{\cdot}5}O_{12}$ below 70°K.[7] However, the variation of susceptibility with H and T is not explainable entirely on the basis of free ions; χ decreases too slowly with increasing T, and σ is not linear in H even at 70°K.

The variation of the observed Curie temperature with substitution of nonmagnetic ions in octahedral positions is well followed by calculated values (Table 1). The agreement obtained is much better than in the case of the first-order approximation employed earlier.[5] The improvement is a result of better information regarding the distribution of nonmagnetic ions provided by the X-ray diffraction studies of the tin-substituted yttrium–iron garnet system and by the improved model developed above.

4. FERRIMAGNETIC SPINELS

The interpretation of the variation of magnetic moment and Curie temperature in ferrites (ferrimagnetic spinels) may be carried out on the same basis as was used for ferrimagnetic garnets. The magnetic properties of ferrites have been extensively interpreted with great success by NÉEL[2]* on a molecular field basis. NÉEL[15] also early recognized the importance of the statistical variation of the number of interactions possible for different magnetic ions in the case of a ferrite in which nonmagnetic ions were introduced. He showed that the variation of the moment of a nickel-zinc ferrite with zinc content could be understood on the basis of the requirement of at least two interactions of a magnetic ion with those of another coördination before it could take part in ferrimagnetism.

However, the mathematical relations were not stated, nor was the Curie temperature variation considered. The relevant equations are therefore briefly presented below.

* For an excellent survey of this work, see GORTER[18].

In the ensuing presentation we make use of the formula unit

$$(Ma_{1-z}M_z)[Mb_{2-y}M_y]O_4,$$

where Ma and Mb are the tetrahedral and octahedral magnetic ions and M is a nonmagnetic ion. As before, we define

$$k_o = y/2,$$
$$k_t = z,$$
$$3k = 2k_o + k_t.$$

In these terms the magnetic moment is given by

$$n_B(k_o,k_t) = 2\mathscr{M}_b(1-k_o)[1-E_o(k_t)] - \mathscr{M}_a(1-k_t)[1-E_t(k_o)]\mu_B,$$

where $\mathscr{M}_a$ and $\mathscr{M}_b$ are the magnetic moments (on the average) of the magnetic ions in tetrahedral and octahedral sites and

$$E_o(k_t) = 6k_t^5 - 5k_t^6,$$
$$E_t(k_o) = 12k_o^{11} - 11k_o^{12}.$$

If $k_o = 0$, as is often assumed in the case of zinc substitution for divalent ions of nickel, manganese, cobalt, and iron in ferrites, we would have a maximum of $n_B(0,k_t)$ when $\partial n_B(0,k_t)/\partial k_t = 0$. For zinc-substituted manganese ferrite a maximum of n_B would occur at $z = 0{\cdot}436$ and would be about 6·58 μ_B, and similarly for other cases. Observed maxima fall between values of 0·4 and 0·5 for z, and are of about the size expected.[18] However, the observed moment falls off more slowly than would be expected from the equation for n_B with increase of z. On the other hand there is reason to believe that not all zinc enters tetrahedral sites for all concentrations.[18] Furthermore, the behaviour of the moment with variation of H and T at low temperature would have to be taken into account.

Calculation of the dependence of Curie temperature on the concentration and distribution of nonmagnetic ions also proceeds in the same manner as before, when all magnetic ions are equivalent. We have, in the case of a spinel,

$$n(k_o,k_t) = 24(1-k_o)[1-E_o(k_t)](1-k_t)[1-E_t(k_o)]$$

and

$$N(k_0,k_t) = 2(1-k_o)[1-E_o(k_t)] + (1-k_t)[1-E_t(k_o)].$$

Accordingly

$$T_C(k_o,k_t) = \frac{3}{24}\frac{n(k_o,k_t)}{N(k_o,k_t)}\,T_C(0,0),$$

where $T_C(0, 0)$ is chosen in accord with the unsubstituted composition.

It is important to recognize that this equation for T_C is valid only when the character of the linkages does not change with substitution, and only their number changes. Consequently, the equation is valid only when all magnetic ions are identical, as in the $ZnFe_2O_4$–$Li_2Fe_5O_8$ system,[18] or when the substituent ion displaces a magnetic ion on the same site, which may seldom be true. In the case of the $(Zn)[Fe_2]O_4$–$(Fe)[FeNi]O_4$ system,[18] Zn^{2+} displaces Fe^{3+} from a tetrahedral to octahedral site, where it displaces Ni^{2+} from the material. Under this circumstance there are varying numbers of (Fe^{3+})–O^{2-}–$[Fe^{3+}]$ and (Fe^{3+})–O^{2-}–$[Ni^{2+}]$ linkages of different strength which must be taken into account in determining the Curie temperature.

In the case of compounds, such as $MgFe_2O_4$[18] and $CuFe_2O_4$,[18] with a constant content (but varying distribution) of nonmagnetic ions, and only one kind of magnetic ion, namely Fe^{3+}, the above equations apply. Substituted manganese–tin spinels containing only Mn^{2+} also fulfill these requirements.[19]

The presence of magnetic ions excluded from participation in ferrimagnetism, except at very low temperatures, again carries with it the question of the behaviour of the magnetic moment and susceptibility arising from these ions at very low temperatures. From the antiferromagnetism observed in $(Zn)[Fe_2]O_4$[20, 21] and $(Cd)[Fe_2]O_4$[21] below 10°K, it is known that octahedral Fe^{3+} ions excluded from interaction with tetrahedral magnetic ions interact antiferromagnetically with other octahedral Fe^{3+} ions.

The numbers of magnetic ions per formula unit totally without interaction with those in another coördination are $2(1-k_o)k_t^6$ in octahedral sites and $1(1-k_t)k_o^{12}$ in tetrahedral sites; the numbers with but a single interaction with an ion in another coördination are $2(1-k_o)(6k_t^5-6k_t^6)$ and $1(1-k_t)(12k_o^{11}-12k_o^{12})$ respectively as before. Each category of excluded ion may be expected to contribute differently to the magnetic moment at very low temperatures. However, at temperatures above, say, 20°K, all probably contribute nearly as free ions. For this reason low-temperature susceptibility data are important to a fuller understanding of ferrimagnetism and superexchange interactions in heavily substituted ferrites.

5. DISCUSSION AND CONCLUSIONS

The recent synthesis of compounds in the yttrium–iron—calcium–iron–tin (YIG—CIS) garnet system[7, 8] has provided an unusually good basis for simultaneous demonstration of several factors important to the understanding of ferrimagnetism. Because the X-ray scattering factor for Sn^{4+} is twice that for Fe^{3+}, it has been possible to obtain accurate knowledge of the distribution of these ions.[7] Furthermore, only one magnetic ion, namely Fe^{3+}, is involved; its moment is subject to negligible variation with coördination number because its ground state is ${}^6S_{5/2}$. For these reasons ambiguity regarding the distribution of ions and their moments is absent.

With the aid of X-ray crystallographic and magnetic data now available for the YIG—CIS system,[7, 8] two concepts valuable to the understanding of ferrimagnetism have been placed on firmer ground: (1) two or more linkages of a magnetic ion in one sublattice are required with magnetic ions in another sublattice in order that the ion may actively participate in ferrimagnetism, and (2) the Curie temperature is proportional to the number of active superexchange linkages per magnetic ion per formula unit.

The contribution to the magnetic properties of a ferrimagnet at very low temperatures, which is made by those ions excluded from ferrimagnetism at higher temperatures, remains to be fully investigated and understood. At intermediate temperatures it is known that the excluded ions contribute a paramagnetic moment much as free ions would; such paramagnetic ions could also contribute to ferrimagnetic-resonance line width. At lower temperatures, however, all magnetic ions must become ordered in orientation, so that the final state will depend upon weak superexchange interactions. It is under these circumstances that the theory of YAFET and KITTEL[17] regarding the magnetic sublattices which develop becomes very important. Studies of low-temperature spontaneous magnetization, specific heat, ferrimagnetic-resonance linewidth and neutron-diffraction data could greatly

clarify the status of the excluded ions and thus reveal further details about ferrimagnetism.

Acknowledgements—The author is deeply indebted to Dr. S. GELLER for his unstinting contribution of time and the results of his work in extensive correspondence, without which this paper could not have been written. The data contributed by Dr. R. M. BOZORTH and his suggestions regarding the manuscript are gratefully acknowledged. The comments of Dr. P. W. ANDERSON, which were valuable in the consideration of the role of coöperative interaction by superexchange in ferrimagnetism, are very much appreciated.

VOL. 13 1960

REFERENCES

1. KRAMERS H. A., *Physica, s'Grav.* **1**, 182 (1934).
2. NÉEL L., *Ann. Phys. Paris* **3**, 137 (1948).
3. ANDERSON P. W., *Phys. Rev.* **79**, 350 (1950); *Phys. Rev.* **115**, 2 (1959).
4. GILLEO M. A., *Phys. Rev.* **109**, 777 (1958).
5. GILLEO M. A. and GELLER S., *Phys. Rev.* **110**, 73 (1958).
6. GELLER S. and GILLEO M. A., *J. Phys. Chem. Solids* **3**, 30 (1957).
7. GELLER S., BOZORTH R. M., GILLEO M. A. and MILLER C. E., *Bull. Amer. Phys. Soc.* **4**, 241 (1959); *J. Phys. Chem. Solids* **12**, 111 (1960).
8. BOZORTH R. M., GELLER S. and MILLER C. E., *Bull. Amer. Phys. Soc.* **4**, 241 (1959.
   BOZORTH R. M. and GELLER S., *J. Phys. Chem. Solids* **11**, 263 (1959).
9. GELLER S., BOZORTH R. M., MILLER C. E. and DAVIS D. D., *J. Phys. Chem. Solids* **13**, 28 (1960).
10. VILLERS G., PAUTHENET R. and LORIERS J., *J. Phys. Radium* **20**, 382 (1959).
11. GELLER S. and MILLER C. E., *Acta Cryst., Camb.* In the press.
12. GILLEO M. A. and GELLER S., *J. Appl. Phys.* **30**, 297S (1959); *J. Phys. Chem. Solids* **10**, 187 (1959).
13. BERTAUT F. and FORRAT F., *C. R. Acad. Sci. Paris* **242**, 382 (1956); **243**, 1219 (1956).
14. TAUBER A., BANKS E. and KEDESDY H. H., *J. Appl. Phys.* **29**, 385 (1958).
15. NÉEL L., *C. R. Acad. Sci. Paris* **230**, 375 (1950).
16. ALÉONARD R., BARBIER J.-C. and PAUTHENET R., *C. R. Acad. Sci. Paris* **242**, 2531 (1956).
17. YAFET Y. and KITTEL C., *Phys. Rev.* **87**, 290 (1952).
18. GORTER E. W., *Philips Res. Rep.* **9**, 295 (1954); **9**, 321 (1954); **9**, 403 (1954).
19. GILLEO M. A. and MITCHELL D. W., *J. Appl. Phys.* **30**, 20S (1959); *J. Phys. Chem. Solids* **10**, 182 (1959).
20. HASTINGS J. M. and CORLISS L. M., *Phys. Rev.* **102**, 1460 (1956).
21. FRIEDBERG S. A. and BURK D. L., *Conference on Magnetism and Magnetic Materials* p. 62. American Institute of Electrical Engineers Publ. T–78 (Oct. 1955).

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 13. pp. 40–55. Printed in Great Britain.

TEMPERATURE DEPENDENT MAGNETIC CONTRIBUTIONS TO THE HIGH FIELD ELASTIC CONSTANTS OF NICKEL AND AN Fe–Ni ALLOY

G. A. ALERS, J. R. NEIGHBOURS* and H. SATO

Scientific Laboratory, Ford Motor Company, Dearborn, Michigan

(*Received* 21 *September* 1959)

Abstract—Since the fundamental interaction between magnetic moments in a ferromagnet makes a contribution to the total energy, there should also be a corresponding contribution to the elastic constants. In order to observe this effect, the elastic constants C_{44}, $\frac{1}{2}(C_{11}-C_{12})$ and $\frac{1}{2}(C_{11}+C_{12}+2C_{44})$ have been measured in nickel and an Fe–30% Ni alloy through their respective Curie temperatures at high enough applied magnetic fields to eliminate the ordinary "ΔE effect" associated with domain wall motion. The intrinsic magnetic interaction which should produce changes in the elastic constants upon passing through the Curie temperature was clearly observed in both materials. The first and the second derivatives of the exchange energy are estimated from these results. Such an interpretation is relatively straightforward for nickel, but the large volume magnetostriction of the Fe–Ni alloy makes the analysis of the data difficult. The measurements on nickel were extended down to 4·2°K while those on the alloy do not go below room temperature because of a possible martensitic transformation of this alloy.

1. INTRODUCTION

As a ferromagnetic substance is cooled through the Curie temperature, changes in the relative alignment of the magnetic moments give rise to a change in those contributions to the total energy which arise from the interatomic magnetic interaction. Since the elastic constants are derivatives of this total energy with respect to strain, an intrinsic magnetic contribution to the elastic constants should appear when a ferromagnetic material is cooled through its Curie temperature. If this effect could be observed, it would then be possible to determine the nature of the magnetic exchange energy much more precisely than by methods previously adopted. Only the first derivative of the exchange energy, with respect to interatomic distance, is obtained by measurement of volume magnetostriction or thermal expansion, whereas both the first and second derivatives may be obtained from elastic constant measurements. A brief explanation of these principles has been given by one of the authors.[1]

In order to find such intrinsic magnetic contributions to the elastic constants, observations must be made in a magnetic field high enough to eliminate the effects of domain boundary motion which give rise to the well known "ΔE effect" but which are not related directly to the contributions of the fundamental magnetic interaction being sought. Furthermore, it is advantageous that the measured elastic constants be related to the derivatives of the energy in as direct a way as is possible. The bulk modulus and the shear constants C_{44} and $\frac{1}{2}(C_{11}-C_{12})$ satisfy this requirement in cubic materials but none of these has been measured over a temperature range which includes the Curie temperature.

Measurements of the temperature dependence of some of the elastic constants over such a wide temperature range have been attempted for some ferromagnetic substances by several authors. For example, measurements on Young's modulus of polycrystalline nickel have been made by QUIMBY

* Present Address: U.S. Naval Postgraduate School, Monterey, California.

and SIEGEL[2] who found that, at magnetic saturation, the change of the modulus at the Curie point is almost negligible. ENGLER[3], on the other hand, measured the Young's modulus of an Fe–42% Ni alloy through the Curie temperature and found a remarkable change at the Curie point even under a strong external magnetic field, where the contribution from the "ΔE effect" should be negligible.

DÖRING[4] has proposed an explanation for these observations by a kind of relaxation mechanism based on changes in the spontaneous magnetization accompanying the changes in the volume which are a consequence of the modulus measurement. Since this effect is based on the volume dependence of spontaneous magnetization which, in turn, is related to the magnitude of the volume magnetostriction, the effect should be large in Fe–42% Ni and small in Ni, which agrees well with the observations. Since it is a relaxation type mechanism, this explanation would always predict a decrease in the modulus from the values extrapolated from high temperatures, irrespective of the sign of the volume magnetostriction. However, scrutiny of QUIMBY's data[2] shows a slight increase of the modulus below the Curie point which is contrary to Döring's prediction. This indicates the possible existence of the intrinsic effect which is expected from the theory mentioned above.

Since the spontaneous magnetization depends primarily on volume, the "Döring contribution" can be minimized if the measurement is made using isovolumic pure shear distortions. For this reason, Young's modulus measurements by themselves are difficult to interpret in terms of an atomic interaction model. The present experiments make use of the pulse echo method to obtain the three elastic constants C_{44}, $\frac{1}{2}(C_{11}-C_{12})$ and $\frac{1}{2}(C_{11}+C_{12}+2C_{44})$ for single crystals of nickel and and Fe–30% Ni alloy[5] through the Curie temperature. The former two constants are measured by the desirable isovolumic shear distortions and therefore, to a first approximation, do not include any "Döring contribution". However, the last constant is measured by a longitudinal wave distortion and therefore this, as well as the bulk modulus which is calculated by combining these three constants, may include a sizable Döring term. Not only the magnitude of the spontaneous magnetization, but also its direction changes due to the deformation which is a necessary part of the elastic constant measurement process, and this gives rise to still another contribution to the measured elastic constants. Since the quantity which must be determined is the magnetic contribution at fixed magnetization, corrections for these extra contributions must be introduced in order to extract the intrinsic contribution.

Nickel was chosen because there has been a question[6] of the sign of the derivative of the exchange interaction with respect to the interatomic distance. It has, however, the disadvantage that the effect sought may be small and the Curie point is high enough to require accurate measurements at temperatures near the upper limit reached by the present techniques. In contrast to nickel, the alloy was chosen because it has a large magnetic moment per atom and its Curie temperature is near room temperature. Unfortunately, it has the disadvantage of having a martensitic phase transformation which sets a lower limit on the temperature range over which measurements can be made. It also has a large volume magnetostriction which produces a large Döring term and introduces other difficulties which complicate the analysis of the data.

EXPERIMENTAL PROCEDURE

Both crystals used in this investigation were prepared by the Bridgman method. The alloy crystal was purchased from the Alloyd Research Corporation and the pure nickel crystal was grown by one of us from electrolytic nickel. The method used to prepare the [110] oriented acoustic specimens has been previously described.[7,8] The resulting lengths were 1·5077 cm for the nickel specimen and 1·9144 cm for the Fe–30 Ni specimen. Chemical analysis of the iron–nickel alloy showed it to be 30 atomic per cent nickel in iron to within 0·1 per cent of nickel. Metallographic examination of the alloy showed it to be single phased. The acoustic measurements were performed at several different temperatures in an applied magnetic field using apparatus similar to that previously described.[9]

For the nickel crystal, the magnetic field was supplied by an A.D.L. electromagnet. Using a $3\frac{1}{2}$ in. gap with 11 in. pole faces, any field from zero to 12·0 kOe could be applied. Below 77°K, the crystal and its associated isothermal enclosure and resistance thermometer were placed in a helium

bath located between the pole faces. The 4·2°K measurements were made with the crystal immersed in the bath; the higher temperature measurements were taken as the crystal slowly warmed. During the time for making a measurement, the temperature changes were less than 1°K. From 77°K to room temperature the crystal was placed in a small furnace in a liquid nitrogen bath and its temperature measured and controlled as described previously.[8,9] Above room temperature the crystal was placed in a furnace surrounded by a water jacket and the entire assembly placed between the pole faces of the electromagnet. Above 77°K the temperature was measured by a copper–constantan thermocouple and controlled to 0·5 degree during the time of making a measurement.

For the alloy crystal, the applied magnetic field was supplied by a 12 in. Varian electromagnet with a $2\frac{1}{2}$ in. gap. The maximum strength of the magnetic field was about 10 kOe. A small water jacketed furnace was made to contain the alloy crystal and to fit between the pole faces of the magnet. Again the crystal temperature was measured with a copper-constantan thermocouple and controlled to 0·5 degree during the time of measurement. Measurements were not made below room temperature because of the phase transformation which occurs near 0°C for this alloy system.

In all cases the applied magnetic field was measured by a rotating coil component Gaussmeter manufactured by the Rawson Electrical Instrument Company. Since for nickel, the direction of the applied field makes only a small change in modulus[10] (less than 0·1 per cent at 10·0 kOe) no special pains were taken to procure any given direction of applied field. For the alloy crystal, space considerations dictated that the applied field be always parallel to the direction of propagation, i.e., [110]. Room temperature tests showed that the alloy crystal measurements were independent of the direction of the applied field.

The pulsed ultrasonic method yields the time of travel of a sound pulse over the length of the specimen. From these observed transit times, the wave velocities and resulting elastic constants were computed using the room temperature length and density as corrected for thermal expansion and magnetostriction. Both room temperature crystal densities, $\rho = 8{\cdot}907$ g cm^{-3} for nickel and $\rho =$ 8·219 g cm^{-3} for Fe–30% Ni were calculated from X-ray lattice constant data. The data of NIX and MCNAIR[11] were used for the thermal expansion of nickel, while that of the alloy was measured in this laboratory. The directly measured elastic constants $C = C_{44}$, $C' = \frac{1}{2}(C_{11}-C_{12})$ and $C_L = \frac{1}{2}(C_{11}+C_{12}+2C_{44})$ were plotted as a function of

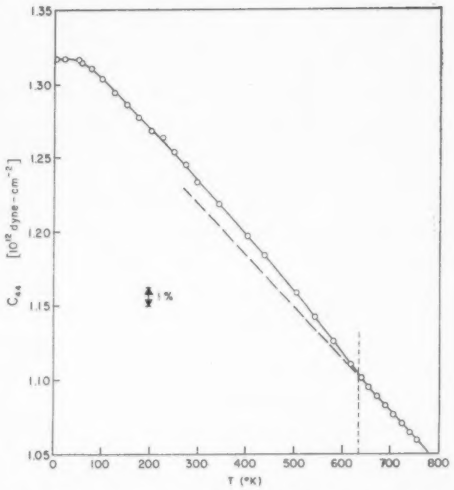


FIG. 1. Temperature variation of the elastic modulus C_{44} of nickel measured at 10 kOe applied field. The vertical dashed line marks the Curie temperature.

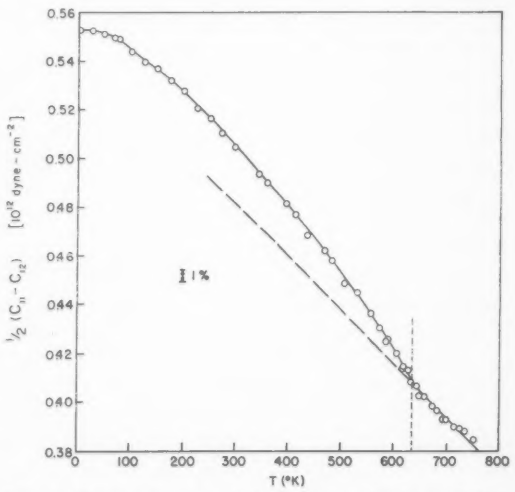


FIG. 2. Temperature variation of the elastic modulus $\frac{1}{2}(C_{11}-C_{12})$ of nickel measured at 10 kOe applied field. The vertical dashed line marks the Curie temperature.

temperature, smooth curves being drawn through the experimental points.

RESULTS FOR NICKEL

The directly measured moduli are shown in Figs. 1–3, where the experimental points, shown as open circles, represent measurements taken at

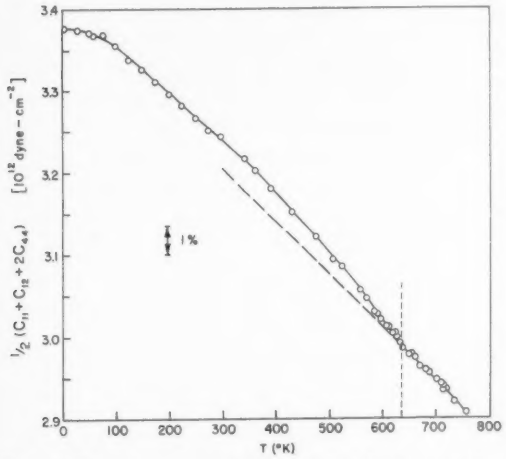


FIG. 3. Temperature variation of the elastic modulus $\frac{1}{2}(C_{11}+C_{12}+2C_{44})$ of nickel measured at 10 kOe applied field. The vertical dashed line marks the Curie temperature.

10·0 kOe applied magnetic field. The dependence of the modulus upon the magnetic field after technical saturation is negligibly small and thus the constants measured at 10·0 kOe are the saturation values characteristic of the temperature of measurement. In these figures the Curie temperature for nickel is indicated by a short vertical dotted line. The sloping dashed lines are extrapolations below the Curie temperature of the essentially linear temperature dependence of the constants measured above the Curie temperature.

Other investigators[7,12–14] using the same techniques have determined the elastic constants of nickel at room temperature and in a somewhat smaller applied magnetic field. Although not plotted in the figures, these previous results are given in Table 1, and disagree with each other and with the results of this paper by about 2 per cent which is more than the conservatively estimated experimental error of 1 per cent. The reasons for this disagreement are not known. Contrary to the observations of LEVY and TRUELL[13] a small "ΔE effect" was observed but not measured quantitatively.

The temperature dependence of the three constants is similar to that observed for other metals; the constants approach zero temperature with zero slope and decrease almost linearly with increasing temperature. The region of near zero slope extends to higher temperatures than those observed for the noble metals. The Debye temperature calculated from this low temperature data is 472°K, which is in agreement with the calorimetric measurements of RAYNE and KEMP[15] if their data is corrected for a magnetic contribution. In passing downward in temperature through the Curie point, the modulus is increased (shown by the experimental data being above the dashed line) by the crystal becoming ferromagnetic. This result agrees with the observations of QUIMBY and SIEGEL[2] and shows that there is indeed a magnetic contribution not describable by Döring's relaxation mechanism.

Since modulus versus pressure data are not available for nickel it is not possible to separate the intrinsic temperature dependence of the constants from that due to thermal expansion. At room

Table 1. *The room temperature saturation values of the elastic constants of nickel as measured by various investigators*

	$\frac{1}{2}(C_{11}+C_{12}+2C_{44})$	C_{44}	$\frac{1}{2}(C_{11}-C_{12})$
Present Work	3·239	1·235	0·504
NEIGHBOURS, BRATTEN and SMITH[7]	3·262	1·238	0·502
BOZORTH, MASON and MCSKIMMIN[12]	3·275	1·22	0·477
LEVY and TRUELL[13]	3·16	1·21	0·475
DE KLERK and MUSGRAVE[14]	3·216	1·247	0·496

temperature, in units of 10^{-4} deg^{-1}, the logarithmic temperature derivatives of the measured constants are:

$$\left(\frac{d\ln C}{dT}\right)_p = -2{\cdot}9, \quad \left(\frac{d\ln C'}{dT}\right)_p = -4{\cdot}5 \times \times \left(\frac{d\ln C_L}{dT}\right)_p = -1{\cdot}9$$

as compared to those above the Curie point which do not include the magnetic contribution which are:

$$\left(\frac{d\ln C}{dT}\right)_p = -3{\cdot}4, \quad \left(\frac{d\ln C'}{dT}\right)_p = -5{\cdot}5 \times \times \left(\frac{d\ln C_{L-}}{dT}\right)_p = -2{\cdot}2$$

Table 2. Values of the elastic constants of nickel at various temperatures as read from smooth curves drawn through the measured data points. The values of C_{11}, C_{12} and the bulk modulus B are calculated from the measured quantities ρV_1^2 ρV_2^2 and ρV_3^2. Applied field = 10 kOe

T (K°)	ρV_1^2 $\frac{1}{2}(C_{11}+C_{12}+2C_{44})$	ρV_2^2 C_{44}	ρV_3^2 $\frac{1}{2}(C_{11}-C_{12})$	C_{11}	C_{12}	B $\frac{1}{3}(C_{11}+2C_{12})$
0	3·377	1·317	0·552	2·612	1·508	1·876
20	3·377	1·317	0·552	2·612	1·508	1·876
40	3·375	1·316	0·551	2·610	1·508	1·875
60	3·370	1·314	0·549	2·605	1·507	1·873
80	3·363	1·309	0·547	2·601	1·507	1·872
100	3·355	1·304	0·544	2·595	1·507	1·870
120	3·344	1·298	0·541	2·587	1·505	1·866
140	3·332	1·291	0·537	2·578	1·504	1·862
160	3·320	1·284	0·533	2·569	1·503	1·858
180	3·309	1·277	0·530	2·562	1·502	1·855
200	3·297	1·270	0·526	2·553	1·501	1·852
220	3·286	1·263	0·522	2·545	1·501	1·849
240	3·274	1·256	0·517	2·535	1·501	1·846
260	3·262	1·249	0·513	2·526	1·500	1·842
280	3·250	1·242	0·508	2·516	1·500	1·839
300	3·239	1·235	0·504	2·508	1·500	1·836
320	3·226	1·228	0·499	2·497	1·499	1·832
340	3·214	1·220	0·494	2·488	1·500	1·829
360	3·201	1·213	0·489	2·477	1·499	1·825
380	3·187	1·205	0·484	2·466	1·498	1·821
400	3·173	1·198	0·479	2·454	1·496	1·815
420	3·159	1·190	0·474	2·443	1·495	1·811
440	3·145	1·183	0·469	2·431	1·493	1·806
460	3·131	1·175	0·463	2·419	1·493	1·802
480	3·116	1·167	0·458	2·407	1·491	1·796
500	3·102	1·159	0·453	2·396	1·490	1·792
520	3·086	1·152	0·447	2·381	1·487	1·785
540	3·071	1·143	0·441	2·369	1·487	1·781
560	3·053	1·135	0·434	2·352	1·484	1·773
580	3·035	1·126	0·428	2·337	1·481	1·766
600	3·018	1·118	0·422	2·322	1·478	1·759
620	3·003	1·109	0·415	2·309	1·479	1·756
640	2·987	1·100	0·408	2·295	1·479	1·751
660	2·973	1·093	0·402	2·282	1·478	1·746
680	2·959	1·086	0·397	2·270	1·476	1·741
700	2·946	1·079	0·394	2·261	1·473	1·736
720	2·932	1·072	0·390	2·250	1·470	1·730
740	2·919	1·065	0·387	2·241	1·467	1·725
760	2·906	1·058	0·384	2·232	1·464	1·720

All these values are very similar to those observed for the noble metals at room temperature. At the Curie temperature the logarithmic derivatives are numerically increased being then

$$\left(\frac{d\ln C}{dT}\right)_p = -3{\cdot}9, \quad \left(\frac{d\ln C'}{dT}\right)_p = -8{\cdot}9 \times$$

$$\times \left(\frac{d\ln C_L}{dT}\right)_p = -2{\cdot}7$$

in the same units.

Table 2 lists readings taken from the smoothed curves drawn to fit the experimental data shown in Figs. 1–3 as well as other constants calculated from these readings. In the table, the absolute accuracy of the measured quantities is certainly better than one-half per cent. The accuracy of the derived constants, especially C_{12} is less, due to propagation of errors.

VOL. 13 1960

RESULTS FOR Fe–30% Ni

First, an auxiliary experiment was undertaken to determine the effect of the orientation of the applied magnetic field with respect to the propagation and polarization directions. In contrast to the results for nickel, no change in velocity was observed as the crystal was rotated about the propagation direction. At 10·0 kOe applied field the maximum effect, if present, was less than 1 part in 5000.

In order to compute the elastic constants from the measured transit times, it is necessary to know the path length and density as a function of temperature. The room temperature density of the alloy crystal was calculated from the X-ray data of OWEN and SULLY[16]. To determine the thermal expansion, a small rod was cut from a portion of the crystal not used in the acoustic measurements. A dilatometer specimen was machined from this rod and after being annealed in vacuum for 4 hr at 800°C, the fractional change in length with temperature was determined in a standard Leitz dilatometer. These results are shown in Fig. 4 where the fractional change in length is plotted as a function of temperature.

An alloy of this composition has a large forced volume magnetostriction. Since the measurements are carried out in applied fields up to 10 kOe, and since the results show a remarkable field dependence, it is necessary to know the value of magnetostriction both for the determination of effects due to volume changes on the elastic constants and for the correction of the path length. The measurement of the magnetostriction as a function of applied field was carried out by an electrical capacity measurement method on the

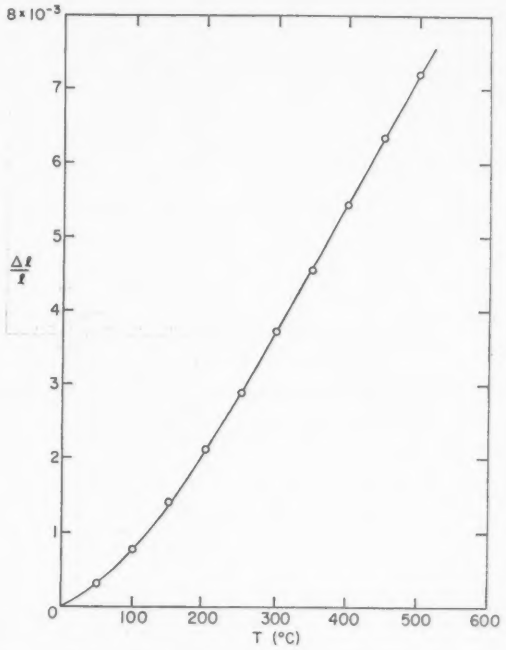


FIG. 4. Relative length change of the Fe–30% Ni alloy as a function of temperature.

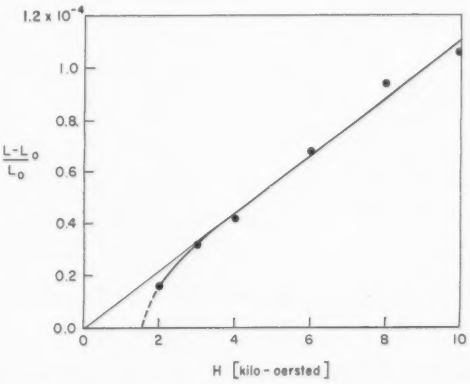


FIG. 5. Relative length change of the Fe–30% Ni alloy as a function of the applied magnetic field at room temperature.

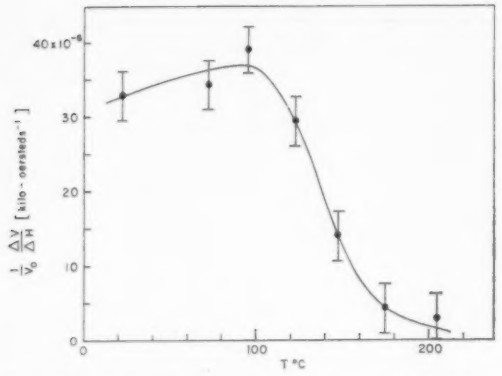


FIG. 6. Forced volume magnetostriction of the Fe–30% Ni alloy as a function of temperature.

crystal specimen with the same physical arrangement as was used for the ultrasonic measurements. Fig. 5 shows the fractional change in length for various values of the applied field at 23°C. Only the part of the forced magnetostriction at high fields is shown and the value obtained of $(1/L_0)(\Delta L/\Delta H) = 11\times10^{-6}$ kOe^{-1} shows good agreement with other values in the literature.[17] The low field magnetostriction is mostly an isovolumic linear magnetostriction and is not shown, since it is not directly related to the intrinsic field dependence of the elastic moduli and because its value is subject to the particular experimental conditions of measurement. However, this part was taken into account for the correction of the path length, even though it was negligible. The values of the forced volume magnetostriction [i.e. $3(1/L_0)(\Delta L/\Delta H)$] obtained at several temperatures are given in Fig. 6. For higher temperatures, this forced magnetostriction becomes less as expected.

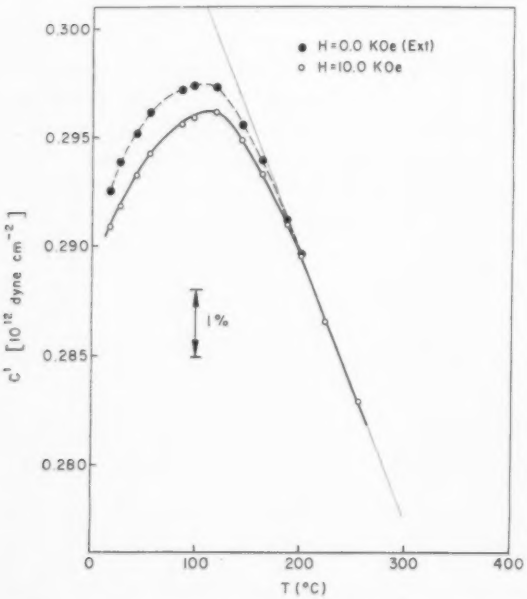


FIG. 7. Variation of the elastic constant $\frac{1}{2}(C_{11}-C_{12})$ of the Fe–30% Ni alloy as a function of temperature. The points at $H = 0{\cdot}0$ kOe were obtained by a linear extrapolation of the high field data to zero applied field.

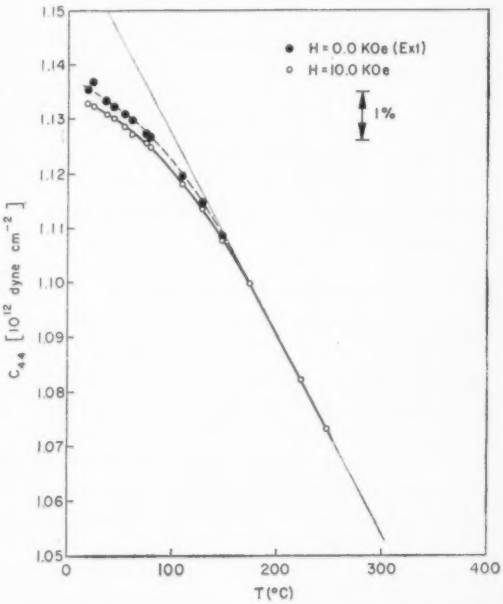


FIG. 8. Variation of the elastic constant C_{44} of the Fe–30% Ni alloy as a function of temperature. The points at $H = 0{\cdot}0$ kOe were obtained by a linear extrapolation of the high field data to zero applied field.

The computed elastic constants C, C' and C_L are plotted in Figs. 7–9 where the open circles represent data obtained at 10·0 kOe and the closed circles represent data obtained by linearly extrapolating measurements of the field dependence of the moduli to zero field. In this way the zero field data may be considered as representing the elastic constants of a single domain crystal in zero

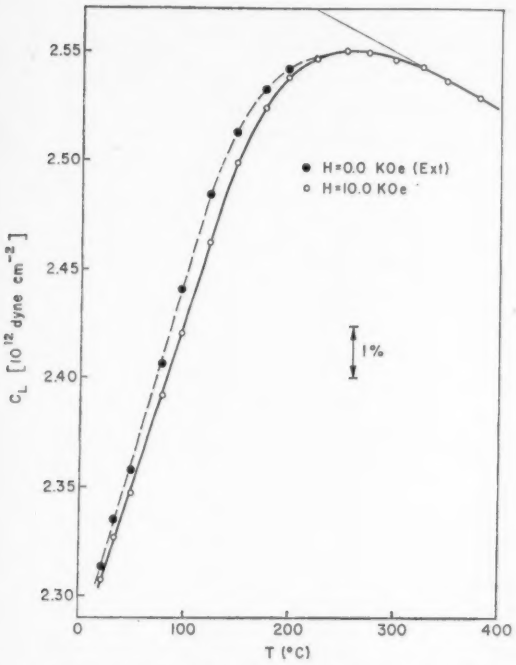


FIG. 9. Variation of the elastic constant $\frac{1}{2}(C_{11}+C_{12}+2C_{44})$ of the Fe–30% Ni alloy as a function of temperature. The points at $H = 0{\cdot}0$ kOe were obtained by a linear extrapolation of the high field data to zero applied field.

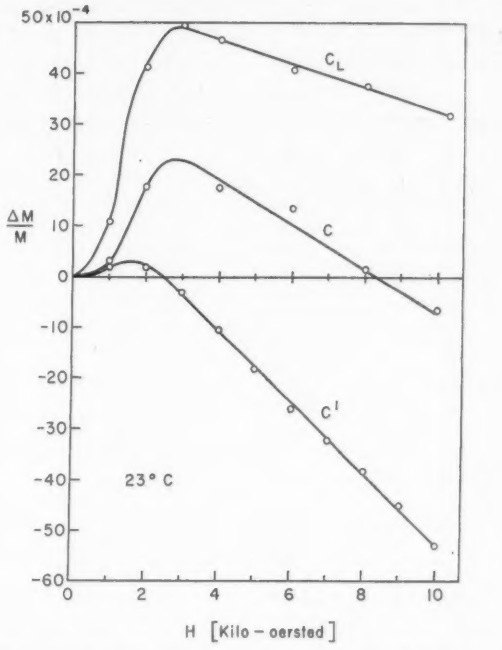


FIG. 10. Relative change of the elastic moduli of the Fe–30% Ni alloy with applied magnetic field at room temperature. The quantity M stands for any of the three moduli C, C' or C_L.

applied field. For all three constants, the low field values are always greater than the higher field values; the difference is largest for C' where it amounts to about $\frac{1}{2}$ per cent at 10 kOe. The two shear moduli C and C' begin to deviate appreciably from the high temperature straight line behavior at about 180°C. Data shown by other investigators[17] of magnetic properties as well as our own data indicate that the Curie temperature of the alloy is in this region.

The temperature dependence of C is similar to that observed for nickel and other metals in that the slope is negative. The two constants C' and C_L are anomalous, since their slopes are initially positive and change to negative with increasing temperature. The transition to a linear dependence with negative slope occurs at a much higher temperature for C_L than for the two shear constants. In all cases, the room temperature elastic constants are much lower than the values extrapolated from the linear region at high temperatures. The deviation amounts to 20 to 30 per cent, indicating a definite contribution of magnetic origin below the Curie point.

Table 3 lists values read from smoothed curves of the temperature dependence of C, C' and C_L shown in Figs. 7–9; the other related constants were calculated from these three. Although not shown in graphs, inspection of the table shows that with decreasing temperature, the elastic anisotropy C/C' passes through a minimum at about 160°C and increases rapidly. Curiously, the deviations from the Cauchy relations $(C_{12}-C_{44})$ change sign at this same temperature. Below about 160° the difference is negative and above it is positive. The bulk modulus, derived principally from C_L, has nearly the same temperature dependence as C_L except that it increases linearly with temperature even above 300°C.

As mentioned previously, increasing the magnetic field changes the observed acoustic wave transit time. Part of this change can be accounted

Table 3. Values of the elastic constants of an iron–30 atomic per cent nickel alloy at various temperatures. The values C_{11}, C_{12} and the bulk modulus B are calculated from the measured quantities ρV_1^2, ρV_2^2, and ρV_3^2. The two shear constants C_{44} and $\frac{1}{2}(C_{11}-C_{12}$ were measured only to 250°C, and the entries in the table for higher temperatures were obtained from a linear extrapolation. Applied field = 10kOe

T (°C)	ρV_1^2 $\frac{1}{2}(C_{11}+C_{12}+2C_{44})$	ρV_2^2 C_{44}	ρV_3^2 $\frac{1}{2}(C_{11}-C_{12})$	C_{11}	C_{12}	B $\frac{1}{3}(C_{11}+2C_{12})$
20	2·304	1·132	0·291	1·463	0·881	1·075
40	2·335	1·130	0·293	1·498	0·912	1·107
60	2·367	1·128	0·295	1·533	0·944	1·141
80	2·397	1·124	0·296	1·569	0·978	1·175
100	2·429	1·120	0·296	1·605	1·013	1·210
110	2·444	1·118	0·296	1·622	1·030	1·227
120	2·460	1·116	0·296	1·640	1·048	1·245
130	2·476	1·113	0·296	1·659	1·067	1·264
140	2·489	1·110	0·295	1·674	1·084	1·280
160	2·512	1·104	0·294	1·701	1·114	1·310
180	2·528	1·098	0·292	1·722	1·139	1·333
200	2·540	1·090	0·289	1·739	1·161	1·353
220	2·546	1·083	0·287	1·751	1·177	1·368
240	2·550	1·076	0·285	1·759	1·189	1·379
260	2·550	1·068	0·282	1·765	1·200	1·388
280	2·549	1·061	0·280	1·769	1·209	1·395
300	2·547	1·053	0·278	1·772	1·216	1·401
320	2·543	1·046	0·276	1·773	1·222	1·406
340	2·539	1·039	0·273	1·774	1·227	1·409
360	2·534	1·031	0·271	1·772	1·230	1·413
380	2·529	1·024	0·269	1·774	1·237	1·416

for by changes in the sample dimensions, but actually this part is negligible compared to the total difference. Thus the change must be interpreted as a change in modulus, probably due in part to the increase in the spontaneous magnetization due to the applied external field and partly due to the volume change of the material through the volume magnetostriction as shown in Fig. 5. In Fig. 10, the fractional change in modulus $\Delta M/M$ is plotted against applied field for all three acoustic modes. Data for this figure were obtained in a manner previously described[10] by observing changes in the time of arrival of a particular echo as a function of applied field. The initial change is greatest for C_L and least for C', while the slopes of the linear portions of the curves are in the reverse order. However, this initial change is mainly due to the redistribution of the magnetic domains (ΔE effect) and is outside the scope of the present study.

As the temperature changes, the fractional change in modulus with field also changes. This is shown in Fig. 11 where the slopes of the linear portions of curves like those of Fig. 10 are plotted against temperature. In this figure no points are shown for the bulk modulus B since it and its change with field are calculated from the experimentally

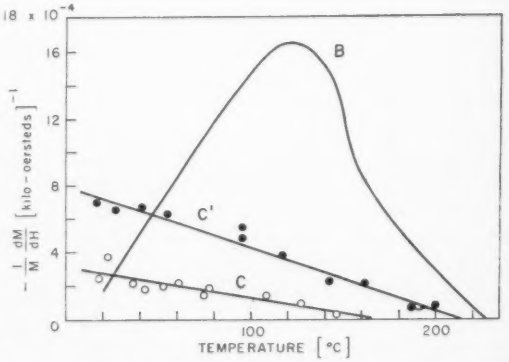


FIG. 11. The field dependence of the moduli of the Fe–30% Ni alloy as a function of temperature. The quantity M stands for any of the three moduli C, C' or B.

determined quantities. The logarithmic derivatives with respect to the field of the shear constants C and C' decrease monotonically with increasing temperature, while that of the bulk modulus passes through a maximum.

INTERPRETATION

The basis for interpretation has already been briefly discussed by SATO[1]. It was shown, based on a localized model of ferromagnetism, that the magnetic contribution to the elastic constants could be broken into a central force term and a non-central force term. The latter has been shown experimentally[10] to be negligibly small. Thus the observed changes below the Curie temperature can be associated with the central force term only. This term arises essentially from the isotropic part of the exchange interaction which, for simplicity, is considered as a short range force so that it is necessary only to consider the nearest neighbors. With these assumptions and using the Ising model, the energy of the ferromagnetic material can be written approximately as:

$$E = E_0 - \tfrac{1}{2}Nz(I/I_0)^2 J(r) \tag{1}$$

where $J(r)$ is the exchange interaction energy, I/I_0 the relative magnetization, z the number of nearest neighbors, N the number of atoms per unit volume and E_0 the collection of all energies not related to the alignment of magnetic moments. From this it follows that for a face centered cubic material, the adiabatic (i.e. fixed magnetization) elastic constants for the equilibrium interatomic distance r_0 are:

$$C_I = C_{44} = C_0 - \frac{N}{2}\left(\frac{I}{I_0}\right)^2\left(r_0^2\frac{d^2J}{dr^2} + 3r_0\frac{dJ}{dr}\right) \tag{2}$$

$$C_I' = \tfrac{1}{2}(C_{11} - C_{12}) = C_0' - \frac{N}{4}\left(\frac{I}{I_0}\right)^2\left(r_0^2\frac{d^2J}{dr^2} + 7r_0\frac{dJ}{dr}\right) \tag{3}$$

$$B = \tfrac{1}{3}(C_{11} + 2C_{12}) = B_0 - \frac{2N}{3}\left(\frac{I}{I_0}\right)^2 r_0^2\frac{d^2J}{dr^2} \tag{4}$$

The terms C_0, C_0' and B_0 represent the values the elastic constants would have if ferromagnetism did not occur. The second terms are the intrinsic magnetic contributions to the elastic constants and exhibit a dependence upon the magnitude of the spontaneous magnetization. Actually, measurements cannot be performed in a truly adiabatic manner since the relative magnetization may be changed by the distortion of the crystal. Thus several corrections must be applied before one obtains the elastic constants at constant magnetization.

To a first approximation the quantities C_H and C_H' measured at constant field are at fixed magnitude of magnetization because they result from isovolumic distortions which do not affect the magnitude of the magnetization. This follows from the known fact that shearing distortions (spontaneous morphic magnetostriction), which result from a change in the spontaneous magnetization, are small compared to the accompanying volume distortions (spontaneous volume magnetostriction). The possibility of magnetization direction changes induced by the measurement process must be taken into consideration. Experimentally,[10] this effect was found to be almost negligible Thus $C_H = C_I$ and $C_H' = C_I'$ for all practical purposes. The measured bulk modulus B_H does involve a volume change and a correction for the effect of the associated magnetization change must be applied to obtain the B_I at fixed magnetization given in equation (6). This correction has been estimated by Döring to be

$$\frac{1}{B_I} = \frac{1}{B_H} - \omega^2 \Big/ \frac{dI}{dH} \tag{5}$$

where ω is the volume magnetostriction dV/dH. It is worthwhile noting that the correction is always such that $B_I > B_H$.

Now the problem is to separate the magnetic contribution from the measured modulus so that the theoretical analysis can be done using equations (2), (3) and (4). By definition, the magnetic contribution is the difference between the elastic constants in the ferromagnetic state and those in the paramagnetic state. This difference can be obtained by taking the difference between the measured values at some temperature below the Curie point and the values extrapolated to that temperature from the paramagnetic region. An examination of the temperature dependence above the Curie point of the elastic constants of ordinary metals as

well as that of ferromagnetic metals indicates that a linear extrapolation is probably quite satisfactory particularly as long as the volume of the material is a linear function of the temperature. Unfortunately, the appearance of spontaneous magnetization produces nonlinear volume changes below the

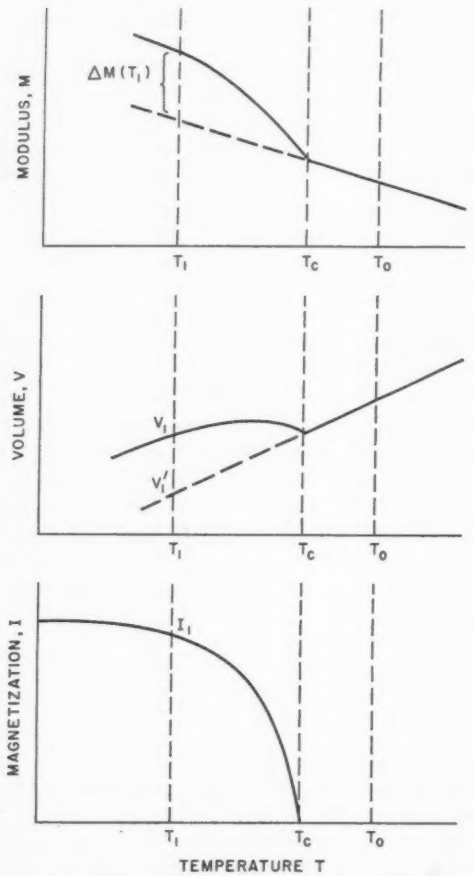


FIG. 12. Schematic representation of the variation of the volume, modulus and magnetization of a ferromagnetic material.

Curie point which give rise to additional changes in the elastic constants. A correction for this effect must be made before the appropriate analysis can be accomplished.

In view of this rather complicated situation, it is convenient to derive some thermodynamic relations among the various measurable quantities in order that the correction terms can be identified and determined. A schematic representation of the variation of the modulus, volume and magnetization of a ferromagnetic material with temperature is shown in Fig. 12. In the figure, the Curie temperature is denoted by T_c and some temperature well below the Curie point by T_1. The dashed lines show a linear extrapolation of the modulus and volume from above the Curie point into the ferromagnetic region. V_1 and I_1 are the actual volume and magnetization at temperature T_1 and V_1', is the extrapolated value of the volume at T_1. The difference between the measured modulus at T_1, and the hypothetical "paramagnetic modulus" at this temperature which is obtained by a linear extrapolation of the high temperature modulus measurements, is $\Delta M(T_1) = M(T_1,\ I_1,\ V_1) - M(T_1,\ 0,\ V_1')$. On the other hand, the intrinsic contribution at T_1 is defined by $M(T_1,\ I_1,\ V_1) - M(T_1,\ 0,\ V_1)$. Thus the modulus difference ΔM is equal to the intrinsic contribution plus a term which corrects for the change in volume from V_1' to V_1.

Let us assume that an elastic modulus M is a function of the volume V, the temperature T and the magnetization I. Here I means the intrinsic magnetization of a ferromagnetic substance and not the magnetization connected with the technical magnetization process. Its selection as a thermodynamic variable is not quite rigorous, but for the present purpose, it seems satisfactory except in the region very near to the Curie point. The magnetization I is itself a function of T, V and the external field H. Under these assumptions, we can easily derive the following thermodynamic relation for the temperature dependence of the measured modulus at constant pressure and magnetic field below the Curie point:

$$\frac{M}{dT} = \left(\frac{\partial M}{\partial T}\right)_{VI} + \left(\frac{\partial M}{\partial V}\right)_{TI}\frac{dV}{dT} + \left(\frac{\partial M}{\partial I}\right)_{TV}\frac{dI}{dT} \quad (6)$$

where $(\partial M/\partial T)_{VI}$ is the intrinsic temperature coefficient of the elastic modulus. Above the Curie point, the magnetization and its derivatives are zero, so that the temperature dependence of M is given by

$$\left(\frac{dM}{dT}\right)_{I=0} = \left(\frac{\partial M}{\partial T}\right)_{0V} + \left(\frac{\partial M}{\partial V}\right)_{0T}\left(\frac{dV}{dT}\right)_{0} \quad (7)$$

where $(dV/dT)_0$ is the specific coefficient of volume expansion for the paramagnetic state. Both dM/dT and $(dM/dT)_0$ can be read from the measurements given in the previous sections. The volume expansion coefficients dV/dT and $(dV/dT)_0$ can be obtained from thermal expansion data and hereafter will be called β and β_0 respectively. In principle, the term dI/dT can also be measured. Integrating equation (6) and (7) from a temperature T_0 above the Curie temperature to T_1, the difference in modulus ΔM can be written,

$$\begin{aligned}\Delta M &= [M(T_1, I_1, V_1) - M(T_1, 0, V_1)] + \\ &\quad + [M(T_1, 0, V_1) - M(T_1, 0, V_1')] \\ &= \left\{\int_0^{I_1}\left(\frac{\partial M}{\partial I}\right)_{TV} dI + \int_{T_0}^{T_1}\left[\left(\frac{\partial M}{\partial V}\right)_{TI} - \left(\frac{\partial M}{\partial V}\right)_{T0}\right]\left(\frac{dV}{dT}\right) dT\right\} + \\ &\quad + \left\{\int_{T_0}^{T_1}\left(\frac{\partial M}{\partial V}\right)_{T0}\left[\left(\frac{dV}{dT}\right) - \left(\frac{dV}{dT}\right)_0\right] dT\right\} \qquad (8)\end{aligned}$$

where the intrinsic temperature dependence on the magnetic exchange terms has been neglected so that the equality

$$\left(\frac{\partial M}{\partial T}\right)_{VI} = \left(\frac{\partial M}{\partial T}\right)_{V0}$$

may be allowed. The first bracketed term on the right is the intrinsic magnetic contribution to the modulus and is equal to the right hand terms in equations (2), (3) and (4) which involve the exchange energy. The second bracketed term on the right is the correction for the change in volume occurring with the appearance of the spontaneous magnetization. This correction can be made if the quantity $(\partial M/\partial V)_{T0}$ and the volume expansion coefficients are known. The latter are easily obtained from existing thermal expansion data. The former can be obtained from measurements of the pressure dependence of the elastic moduli above the Curie point. This can be done in principle but was not carried out in the present case. Instead, it was possible to make estimates of the corrections making use of the field dependence data and the thermal expansion data. A detailed discussion of these estimates is given in the Appendix.

In the case of nickel, the magnetic field dependence of the modulus was observed to be too small to measure. However, examination of the thermal expansion data of Nix and McNair[11] shows that, by a fortuitous circumstance, the integral of

$$\left[\left(\frac{dV}{dT}\right) - \left(\frac{dV}{dT}\right)_0\right]$$

in equation (8) vanishes at 480°K. Thus by making the reasonable assumption that $(\partial M/\partial V)_0$ varies only slightly with temperature between 480°K and 700°K, the correction term is negligible at this temperature and the intrinsic contribution may be determined.

The values for the derivatives of the exchange energy thus obtained are:

$$\tfrac{1}{2} Nr_0 \frac{dJ}{dr} = (-11 \pm 1) \times 10^9 \text{ dyne cm}^{-2}$$

$$\tfrac{1}{2} Nr_0^2 \frac{d^2J}{dr^2} = (+12 \pm 5) \times 10^9 \text{ dyne cm}^{-2}$$

The errors indicated arise from the uncertainty in making the extrapolation from 700°K. The values given above define the derivatives of J at the interatomic spacing r_0 which exists at 480°K. The sign of the first derivative agrees with a recent measurement of the volume magnetostriction,[6] and is also consistent with the generally accepted information concerning the location of this element with respect to the "Bethe-Slater curve" of the exchange energy. This data also represents the first information concerning the second derivative of the exchange energy.

Applying these results on the derivatives of J to equation (4) it is predicted that the intrinsic contribution to the bulk modulus of nickel is given by $B_I = B_0 - (9+4) \times 10^9$ dyne/cm^2. However, the observation gives $B_H = B_0 + 6 \times 10^9$ dyne/cm^2. Correction for the Döring term makes this disagreement even worse since equation (5) shows $B_I > B_H$. This lack of consistency implies that some other contributions might have been overlooked or that the assumption of the strictly localized model is inadequate. This point will be discussed later.

For the iron–nickel alloy, the correction term can never be neglected. Again pressure data are not available, and the fortuitous intersection of the extrapolated volume versus temperature dependence curve with the actual data, as in the case

of pure nickel, does not occur. However, the measured magnetic field dependence of the moduli provides an additional relationship which can be used to eliminate the $(\partial M/\partial V)_0$ term in equation (8). At constant temperature and pressure the magnetic field dependence of the modulus is

$$\frac{dM}{dH}=\left(\frac{\partial M}{\partial I}\right)_{TV}\frac{dI}{dH}+\left(\frac{\partial M}{\partial V}\right)_{IT}\frac{dV}{dH} \qquad (9)$$

where dM/dH and dV/dH are measured quantities. dI/dH can also be measured in principle, but it was not carried out in the present case. Equations (9) and (6) (temperature dependence) are composed of essentially the same terms if they are integrated with respect to either H or T. However, the expression for the temperature dependence includes an extra term relating to the ordinary thermal expansion. The $(\partial M/\partial I)_{TV}$ term involves the exchange integral derivatives which we seek but dI/dH and $(\partial M/\partial V)_{IT}$ represent two new unknowns. Therefore, by making use of the temperature dependence of the moduli (equation 6) over a small temperature interval near room temperature, it is possible to eliminate $(\partial M/\partial V)_{IT}$. Rather restrictive limits may be placed on dI/dH by placing reasonable limits on $(\partial C/\partial V)_0$ and $(\partial C'/\partial V)_0$ and making use of the fact that the same value of dI/dH must fit the measurements on both C and C'. A detailed description of applying the above procedure to the alloy measurements is given in the Appendix where it is shown that the exchange energy derivatives fall within the limits,

$$+9\times10^9>\frac{N}{2}r_0\frac{dJ}{dr}>+5\times10^9 \text{ dyne/cm}^2$$

$$-6\times10^9>\frac{N}{2}r_0^2\frac{d^2J}{dr^2}>-25\times10^9 \text{ dyne/cm}^2$$

These establish that the first derivative is positive and the second negative which is consistent with the generally accepted ideas concerning the location of this material with respect to the "Bethe-Slater" curve.

The intrinsic contribution to the bulk modulus can be calculated from these results and gives

$$B_I=B_0+(20\pm12)\times10^9 \text{ dyne/cm}^2$$

which would be consistent with the measurements if an appropriate correction for the Döring term and for the volume dependence of the bulk modulus could be made.

SUMMARY

The high field elastic constants C, C' and $\frac{1}{2}(C_{11}+C_{12}+2C_{44})$ were measured over a wide temperature range which included the Curie point for a nickel and an Fe–30% Ni alloy single crystal with the purpose of obtaining detailed information about the exchange energy. In both materials, clearly visible changes in the elastic constants were observed at the Curie point and below. For nickel, this change was an increase of about 3 per cent above the value extrapolated from the paramagnetic state. In the case of the alloy, the change was a decrease which is so large that even the temperature coefficients change sign. The change in elastic constants due to the appearance of spontaneous magnetization was extracted from the measurements and was interpreted in terms of a theory based on a localized model of ferromagnetism. Although an exact knowledge of the wave functions is necessary to calculate the elastic constants, the problem was simplified by making the approximation of a strictly localized model in which a magnetic moment is assumed to be placed on each lattice point.

For evaluating the intrinsic magnetic contribution, a correction had to be made for the change in volume accompanying the appearance of the spontaneous magnetization. For nickel, this correction was not very large and the interpretation was relatively straightforward. However, in the case of the alloy, this contribution amounts to a sizable effect but the correction could still be made by utilizing the results of the field dependence of the elastic constants. Based on the measurements of the shear constants alone the following values were obtained for the derivatives of the exchange energy in units of 10^9 dyne/cm^2:

Nickel:

$$\frac{N}{2}r_0\frac{dJ}{dr}=-11\pm1 \qquad \frac{N}{2}r_0^2\frac{d^2J}{dr^2}=+12\pm5$$

Fe–30% Ni:

$$\frac{N}{2}r_0\frac{dJ}{dr}=+7\pm2 \qquad \frac{N}{2}r_0^2\frac{d^2J}{dr^2}=-15\pm9$$

The errors quoted arise mainly from extrapolation

uncertainties in nickel while those for the alloy come from the lack of a knowledge of the volume dependence of its elastic constants. For nickel, the first derivative of the exchange energy with respect to the atomic distance is negative and the second derivative positive. On the other hand, in the case of the alloy, the first derivative is positive and the second derivative negative. These results for the first derivatives of both materials agree with the generally accepted knowledge obtained from the measurement of the volume magnetostriction, thermal expansion, etc. The second derivative of the exchange energy has never before been obtained.

VOL. 13 1960

Although there may be some uncertainties about the absolute values because of the various corrections, the signs are not subject to small changes in the corrections and thus rather clear-cut information concerning the location of these metals on the "Bethe-Slater curve" for the exchange energy is obtained, although this does not mean that each element or alloy should lie on a universal curve. These results are not inconsistent with the generally accepted ideas based on other measurements.

The values derived from the shear constants were also used to calculate the intrinsic magnetic contribution to the bulk modulus. In the case of nickel this intrinsic contribution is negative. The relaxation effects, as discussed by Döring also make a negative contribution. However, the observed magnetic contribution to the bulk modulus is positive and this represents a real inconsistency in the interpretation. The cause of this difference remains unexplained. In the case of the alloy, the intrinsic contribution to the bulk modulus is positive, instead of being negative as observed. However, this result may not be inconsistent if one includes the relaxation effects.

The failure of the present model to give the correct contribution to the bulk modulus may indicate the failure of the basic assumption of a strictly localized model of ferromagnetism. It is well known, especially in the case of nickel, that such a model fails. As a matter of fact, in this metal at least two different electronic structures $3d^9$ and $3d^{10}$ are often considered to be mixed. In a collective electron model, on the other hand, the energy of ferromagnetism depends on the volume of the crystal as well as on the degree of magnetization, and, therefore, an extra term in the bulk modulus is to be expected in going from the paramagnetic state to the ferromagnetic state. However, this model by itself cannot explain a change in the shear moduli because these involve isovolumic distortions which are not influenced by energy terms which depend on the volume only. The true situation is probably something in between the localized and collective electron models of ferromagnetism. This discrepancy in the bulk modulus might be an indication of the deviations from a strictly localized model.

Acknowledgements—The authors are indebted to many members of the Scientific Laboratory staff for their help during these experiments. Particularly, Mr. R. L. FORGACS of the Electrical Department who designed and built a greatly improved system for making the ultrasonic measurements which was used for the nickel data. Mr. F. J. SEMER of the Metallurgy Department performed the thermal expansion measurements on the Fe–Ni alloy specimen and Mr. D. WALDORF of the Physics Department assisted in making many of the acoustic measurements.

REFERENCES

1. SATO H. *J. Appl. Phys.* **29**, 456 (1958); *Phys. Rev.* **109**, 802 (1958).
2. SIEGEL S. and QUIMBY S. L. *Phys. Rev.* **49**, 931 (1936).
3. ENGLER O. *Ann. Phys.* **31**, 145 (1938).
4. DÖRING W. *Ann. Phys.* **32**, 465 (1938).
5. NEIGHBOURS J. R., ALERS G. A. and SATO H. *Bull. Amer. Phys. Soc.* **3**, 123 (1958).
6. SNOEK J. L. *Physica, s'Grav.* **4**, 853 (1937).
   AZUMI K. and GOLDMAN J. E. *Phys. Rev.* (1955).
7. NEIGHBOURS J. R., BRATTEN F. W. and SMITH C. S. *J. Appl. Phys.* **23**, 389 (1952).
8. NEIGHBOURS J. R. and ALERS G. A. *Phys. Rev.* **111**, 707 (1958).
9. ALERS G. A. and NEIGHBOURS J. R. *J. Phys. Chem. Solids* **7**, 58 (1958).
10. ALERS G. A., NEIGHBOURS J. R. and SATO H. *J. Phys. Chem. Solids* **9**, 21 (1958).
11. NIX F. C. and MCNAIR D. *Phys. Rev.* **60**, 567 (1941).
12. BOZORTH R. M., MASON W. P. and MCSKIMMIN H. J. *Bell. Syst. Tech. J.* **30**, 970 (1951).
13. LEVY S. and TRUELL R. *Rev. Mod. Phys.* **25**, 140 (1953).
14. DE KLERK J. and MUSGRAVE M. J. P. *Proc. Phys. Soc. Lond.* B**68**, 81 (1955) and DE KLERK J. *Proc. Phys. Soc. Lond.* **73**, 337 (1959).
15. RAYNE J. and KEMP W. R. G. *Phil. Mag.* **1**, 918 (1956).
16. OWEN E. A. and SULLY A. H. *Phil. Mag.* **31**, 314 (1941).
17. BOZORTH R. M. *Ferromagnetism.* Van Nostrand, New York (1950).
18. WEISS P. and FORRER R. *Ann. Phys.* **5**, 153 (1926).

APPENDIX

The intrinsic magnetic contribution to the elastic constants at a temperature T_1 is defined by equation (8) in the text to be

$$\int_0^{I_1}\left(\frac{\partial M}{\partial I}\right)_{TV} dI+\int_{T_0}^{T_1}\beta\left[\left(\frac{\partial M}{\partial V}\right)_{TI}-\left(\frac{\partial M}{\partial V}\right)_{T0}\right] dT$$

$$=\Delta M-\left(\frac{\partial M}{\partial V}\right)_{T0}\int_{T_0}^{T_1}[\beta-\beta_0]\,dT \quad \text{(A.1)}$$

For clarity, this will be rewritten and discussed in terms of the elastic constant C. Exactly parallel arguments can be made for the constant C'. In terms of C, equation (A.1) becomes

$$-\tfrac{1}{2}N(I/I_0)^2\left[r_0^2\frac{d^2J}{dr^2}+3r_0\frac{dJ}{dr}\right]$$

$$=\Delta C-\left(\frac{\partial C}{\partial V}\right)_{T0}\int_{T_0}^{T_1}[\beta-\beta_0]\,dT \quad \text{(A.2)}$$

Here β is the specific volume expansivity and ΔC is the difference between the measured modulus at T_1 and the modulus obtained from a linear extrapolation of the elastic constant data from above the Curie point down to T_1. Such a linear extrapolation is probably quite good in any case, and especially so, as long as the volume is a linear function of temperature. Actually the thermal expansion of a normal metal is not strictly linear, but since we assume a linear modulus variation it is consistent to set the "paramagnetic" volume expansion coefficient β_0 equal to a constant. In this way the true expansion coefficient β below the Curie point will include not only the non-linear effects introduced by the appearance of spontaneous magnetization but also all the non-linear terms common to all normal metals. From this definition, it follows that the integral of $(\beta-\beta_0)$ in equation (A.2) is simply the difference between the true volume at T_1 and the volume obtained by a linear extrapolation of the volume versus temperature curve from the range above the Curie point down to T_1.

In the case of nickel, examination of the thermal expansion data of NIX and MCNAIR[11] shows that near 480°K the true volume equals the linearly extrapolated value. Below this temperature the true volume is larger than the extrapolated value while between it and the Curie point the true volume is less than the extrapolation. Thus at 480°K the magnetic contributions to the volume just cancel the contributions from the non-linear variation of the volume associated with the non-magnetic state. This fortuitous circumstance makes the integral of $(\beta-\beta_0)$ equal to zero at 480°K and no correction term is involved. Using the subscript x to denote the extrapolated values, the differences between the extrapolated and observed shear constants C and C' can be read from Fig. 1 and 2 and are found to be

$$C(T_1)-C_x(T_1)=1{\cdot}22\pm0{\cdot}13\times10^{10}\ \text{dyne cm}^{-2}$$

$$C'(T_1)-C_x'(T_1)=1{\cdot}91\pm0{\cdot}21\times10^{10}\ \text{dyne cm}^{-2}$$

at the temperature $T_1=480$°K. The value of relative magnetization $I/I_0=0{\cdot}782$ was taken from the measurements of WEISS and FORRER[18] at 480°K. Since this temperature is well below the Curie temperature the approximations used in deriving equation (1) are valid. Using these shear constant differences and the value of relative magnetization, the values for $(N/2)r_0(dJ/dr)$ and $(N/2)r_0^2(d^2J/dr^2)$ given in the text were obtained. Although the temperature dependence of J is negligible these values may change considerably with temperature depending upon the location of r_0 on the "Bethe–Slater" curve.

Considering the Fe–30% Ni alloy, reference to Figure 4 shows that the integral of $(\beta-\beta_0)$ in equation (A.1) is definitely never negligible. Thus a knowledge of $(\partial M/\partial V)_0$ is very important. An independent equation involving $(\partial M/\partial V)_0$ can be obtained from the magnetic field dependence of the elastic constants which was measured for this alloy. From equation (9) in the text, the change in modulus upon application of a 10 kOe magnetic field at a constant temperature is given by

$$\Delta M(H)=\left(\frac{\partial M}{\partial I}\right)_V\Delta I(H)+\left(\frac{\partial M}{\partial V}\right)_I\Delta V(H) \quad \text{(A.3)}$$

where $\Delta I(H)$ and $\Delta V(H)$ are the magnetization change and volume change resulting from this change in field. In terms of the elastic constant C and the derivatives of J, equation (A.3) can be written

$$\Delta C(H)=-2\frac{\Delta I(H)}{I_0}\frac{N}{2}(I/I_0)^2\left(r_0^2\frac{d^2J}{dr^2}+3r_0\frac{dJ}{dr}\right)+$$

$$+\left[\left(\frac{\partial C}{\partial V}\right)_I-\left(\frac{\partial C}{\partial V}\right)_0\right]\Delta V(H)+\left(\frac{\partial C}{\partial V}\right)_0\Delta V(H) \quad \text{(A.4)}$$

This second relation also includes two unknowns,

$$\frac{\Delta I(H)}{I_0}\ \text{and}\ \left[\left(\frac{\partial C}{\partial V}\right)_I-\left(\frac{\partial C}{\partial V}\right)_0'\right].$$

To find these, consider a small change in temperature. There is a magnetization change $\Delta I(T)$ and a volume change $\beta\Delta T$, so that equation (6) in the text can be written

$$\Delta C(T)=\left(\frac{\partial C}{\partial T}\right)_{VI}\Delta T+\left(\frac{\partial C}{\partial V}\right)_{IT}\beta\Delta T+$$

$$+\left(\frac{\partial C}{\partial I}\right)_{TV}\Delta I(T) \quad \text{(A.5)}$$

The quantity $\beta\Delta T$ is the total volume change occurring as the temperature is changed by ΔT and is made up of two contributions as can be seen from the defining relationship,

$$\beta=\frac{dV}{dT}=\left(\frac{\partial V}{\partial T}\right)_{IH}+\left(\frac{\partial V}{\partial I}\right)_{TH}\frac{dI}{dT} \quad \text{(A.6)}$$

The second term describes the change in volume caused by a change in magnetization. Therefore, choosing the temperature interval ΔT such that

$$\int_{\Delta T}\left(\frac{\partial V}{\partial I}\right)_{TH}\frac{dI}{dT}\,dT=\Delta V(H) \qquad \text{(A.7)}$$

we can make $\Delta I(T)=\Delta I(H)$. Under this condition

$$\begin{aligned}\Delta C(T)-\Delta C(H)&=\left(\frac{\partial C}{\partial T}\right)_{VI}\Delta T+\\&+\left(\frac{\partial C}{\partial V}\right)_{IT}\left(\frac{\partial V}{\partial T}\right)_{IH}\Delta T\\&=\left[\left(\frac{\partial C}{\partial T}\right)_{VI}+\left(\frac{\partial C}{\partial V}\right)_{0}\beta_0\right]\Delta T+\\&+\left[\left(\frac{\partial C}{\partial V}\right)_{IT}\left(\frac{\partial V}{\partial T}\right)_{IH}-\left(\frac{\partial C}{\partial V}\right)_{0}\left(\frac{\partial V}{\partial T}\right)_{0}\right]\Delta T\end{aligned} \qquad \text{(A.8)}$$

The first term is the modulus change which would be associated with ΔT above the Curie temperature. By making the reasonable approximation that

$$\left(\frac{\partial V}{\partial T}\right)_{IH}=\left(\frac{\partial V}{\partial T}\right)_{0}=\beta_0,$$

the second term gives the difference

$$\left[\left(\frac{\partial C}{\partial V}\right)_{I}-\left(\frac{\partial C}{\partial V}\right)_{0}\right]$$

which is needed. In order to estimate $\Delta I(H)$, use can be made of the fact that the same $\Delta I(H)$ must satisfy equations for both the C and C' modes. Furthermore, $(\partial C/\partial V)_0$ and $(\partial C'/\partial V)_0$ are restricted in value by the observed negative temperature dependence of C and C' above the Curie point. This temperature dependence is given formally by

$$\frac{dM}{dT}=\left(\frac{\partial M}{\partial T}\right)_{V0}+\left(\frac{\partial M}{\partial V}\right)_{T0}\beta_0 \qquad \text{(A.9)}$$

For most metals the intrinsic part, $(\partial M/\partial T)_{V0}$, is negative, so that one extreme value of $(\partial M/\partial V)_{T0}$ is obtained from setting $(\partial M/\partial T)_{V0}=0$ and associating all the observed temperature dependence with the thermal expansion. The other extreme is to set $(\partial M/\partial V)_{T0}=0$ in which case the observed temperature dependence would arise entirely from the intrinsic term. Imposing the limits $0\geqslant(\partial C/\partial V)_0\geqslant-6{\cdot}9\times10^{12}$ dyne/cm^2 and $0\geqslant(\partial C'/\partial V)_0\geqslant-2{\cdot}2\times10^{12}$ dyne/cm^2 on equation (A.4) for C and C', a solution can be obtained for both only if $0{\cdot}11\geqslant(2\Delta I/I_0)\geqslant0{\cdot}10$. Such an order of magnitude for the relative magnetization change induced by the application of 10 kOe field is rather large compared to most magnetic materials but the data available in the literature[17] on this particular alloy system indicates that such a value is reasonable.

Imposing these limits on $2\Delta I/I_0$ and taking the Curie point to be 190°C so that a value of $(I/I_0)=0{\cdot}88$ could be obtained at room temperature from the Brillouin function for $j=\frac{1}{2}$, the values

$$\tfrac{1}{2}Nr_0\frac{dJ}{dr}=(+7\pm2)\times10^9\ \text{dyne/cm}^2$$

$$\tfrac{1}{2}Nr_0^2\frac{d^2J}{dr^2}=(-15\pm9)\times10^9\ \text{dyne/cm}^2$$

were obtained. The errors indicated arise from the limits on $2\Delta I/I_0$ and although they are large, they do not affect the signs of the derivatives. Because of the long and rather elaborate techniques involved in extracting these results from the data, it is felt that there may be considerable error in the numerical magnitudes of the derivatives of the exchange energy, but the signs are probably correct.

VOL. 13 1960

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 13. pp. 56–64. Printed in Great Britain.

STUDY OF THE CRYSTALLINE TRANSFORMATIONS IN ZnS:Cu, ZnS:Ag AND ZnS:Cu, Al

M. AVEN* and J. A. PARODI

Lamp Development Department, General Electric Company, Cleveland, Ohio

(*Received* 16 *June* 1959; *revised* 14 *September* 1959)

Abstract—An EPR study has shown that the formation of a separate phase of copper (or silver) sulfide in polycrystalline ZnS:Cu, ZnS:Ag and ZnS:Cu, Al triggers the rapid transformation of hexagonal to cubic zinc sulfide. It is believed that, after firing, cuprous sulfide precipitates from solution in hexagonal zinc sulfide and in so doing supplies the energy necessary for nucleation of cubic zinc sulfide. Once nucleation has begun, growth of the cubic phase will be self-sustaining, because energy is released in going from the hexagonal to the cubic structure. Supporting data for the postulated mechanism of nucleation and growth is provided by X-ray diffraction data on the effect of quenching and annealing on the crystal structure of ZnS:9×10^{-4} Cu phosphors. It was found that the optimum nucleation temperature lies close to 700°C and that at this temperature the growth of the cubic phase is reasonably fast.

1. INTRODUCTION

THE existence of two crystalline structures of zinc sulfide, sphalerite and wurtzite, is well known. The former is face-centred cubic and the latter hexagonal close-packed. The two forms are enantiotropic, with a transition point of approximately 1020°C[1] between the low-temperature cubic and high-temperature hexagonal modifications. Within the last decade several new crystalline modifications of zinc sulfide have been reported.[2–4] Each of these is either a hexagonal polytype[3] (e.g. 4*H* and 6*H*), of which 2*H* or wurtzitic ZnS is the prototype, or a rhombohedral polytype (e.g. 15*R*), of which the 3*R* rhombohedral structure reported by BUCK and STROCK[4] is the prototype. A number of investigators have studied the effect of crushing on the hexagonal–cubic phase transition. SCHLEEDE and GANTZCHOW[5], FULLER[6], BRIDGMAN[7], FREY[8], KULLERUD[9], STROCK[10] and VAN WIERINGEN[11] all report that the hexagonal modification inverts to cubic when crushed. SMITH and HILL[12], however, present data which indicates that crushing randomizes the structure and that from both cubic and hexagonal starting materials a similar product is obtained. ALLEN and CRENSHAW[13] have found that iron dissolved in zinc sulfide lowers the transition temperature. KRÖGER[14] reports a similar effect for solid solutions of MnS and CdS in ZnS.

The present work is an attempt to clarify the role played by copper, silver and aluminum in the transformation of zinc sulfide from the hexagonal to the cubic modification. Studies of phase transformations in zinc sulfide have been carried out in the past almost entirely by X-ray diffraction techniques. Recently, however, VAN WIERINGEN[11] showed how electronic paramagnetic resonance could be used to study the hexagonal–cubic phase transition in zinc sulfide. He investigated the effect of pressure on the crystal transformation by observing changes produced in the EPR (electronic paramagnetic resonance) spectrum of 0·05 per cent divalent manganese which had been added to the zinc sulfide prior to firing.† He observed that as the pressure on hexagonal zinc sulfide is increased up to 3500 kg/cm² the EPR spectrum

* Now at General Electric Research Laboratory, Schenectady, N.Y.

† Other experimenters who have reported on the EPR of ZnS:Mn powders are HERSHBERGER and LEIFER[15] and SCHNEIDER and ENGLAND[16]. A detailed EPR analysis of a single crystal of cubic ZnS:Mn is reported by MATARRESE and KIKUCHI[17] and a similar analysis of a single crystal of hexagonal ZnS:Mn is reported by KELLER, *et al.*[18]

gradually changes from one of more than 30 lines to one of six equally spaced lines, which is characteristic of Mn^{++} in a cubic field. Because of the marked change in the Mn^{++} spectrum when its environment changes from that of hexagonal ZnS to cubic ZnS, the method of EPR using Mn^{++} as an indicator was chosen to implement the present investigation of the effect of copper on the hexagonal–cubic phase transition in zinc sulfide.

2. EXPERIMENTAL METHOD

(a) *Preparation of samples*

The first series of samples was prepared from unprefired luminescent-grade zinc sulfide. Manganese, silver, copper and aluminum were added as aqueous solutions of the corresponding nitrate. After mixing, the 10– to 15–g slurries were dried at 100°C and sieved. The samples were fired in fused silica tubes at 1100°C in purified hydrogen sulfide.[19] Comparison of samples fired for 1 and 16 hr at 1100°C showed that firing for 1 hr was sufficient to develop the hexagonal crystal structure and that there was no significant difference in the effect of copper on the hexagonal → cubic transformation between the two firing schedules. Therefore, the 1-hr firing schedule was adopted for all samples. After firing, the samples were removed from the furnace and allowed to cool in air. They were not sieved in order to avoid possible hexagonal → cubic phase transition induced by the mechanical energy of the sieving process; nor were they washed with cyanide to remove excess copper sulfide for the same reason. The manganese concentration was adjusted to 10^{-4} atomic fraction Mn^{++}. This was sufficiently dilute to avoid spin–spin broadening, yet was enough to give a strong signal.

The second series of samples, used to investigate the effect of quenching and annealing on the crystal structure, was prepared by adding 9×10^{-4} g-atoms copper (as nitrate) per mole of zinc sulfide, which had been purified in hydrogen sulfide at 900°C for 1 hr. As before the slurry was dried at 100°C. Ten– to 15–g portions of the dry mixture were fired in fused silica tubes for $\frac{1}{2}$ hr at 1100°C in purified H_2S. After this period the temperature of one group of samples was lowered to the annealing temperature and kept there for 4 hr. A second group of samples was quenched from 1100°C. to room temperature at an average rate of approximately 10°C/sec, kept at room temperature for 10 min and then annealed for $3\frac{1}{2}$ hr. The annealing temperatures for both groups were 500, 700 and 1000°C. After annealing, the samples were screened, washed with a solution of sodium cyanide and subjected to X-ray analysis.

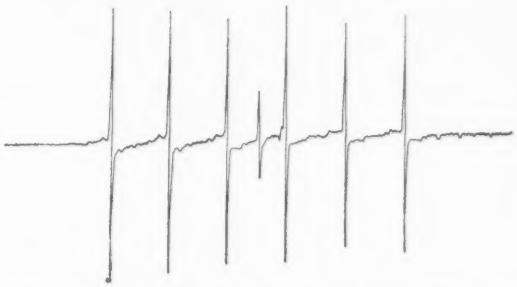


FIG. 1. EPR spectrum of Mn^{++} in polycrystalline cubic zinc sulfide. Prepared by firing ZnS:10^{-4} Mn for 16 hr. at 900°C in H_2S. Central line is resonance of 1,1-diphenyl –2 picryl hydrazyl, $g = 2{\cdot}0036 \pm 0{\cdot}0002$. Magnetic field increases from left to right.

(b) *Electronic paramagnetic resonance*

The zinc sulfide powder was placed in a quartz sample tube 4 mm outside dia. × 3 mm inside dia. × 4 in. long. Roughly $\frac{3}{4}$ in. of sample at the end of the tube was used. This was inserted into the resonant cavity of a Varian Associates Model

FIG. 2. EPR spectrum of Mn^{++} in polycrystalline hexagonal zinc sulfide. Prepared by firing ZnS:10^{-4} Mn for 16 hr at 1100°C in H_2S. Magnetic field increases from left to right.

V–4500 Electronic Paramagnetic Resonance Spectrometer. The resonant cavity operates in the TE_{012} made at a frequency of about 9480 Mc/s. The experimental record, because of ease of instrumentation, is a potentiometric recorder tracing of the first derivative of absorption of microwave energy versus the magnetic field. The reader who is unfamiliar with the theory and practice of EPR

is referred to the reviews of BLEANY and STEVENS[20] BOWERS and OWEN[21] and BAGGULEY and OWEN[22] and to the books of INGRAM[23] and GORDY *et al.*[24].

3. EXPERIMENTAL RESULTS

EPR spectra* of the zinc sulfide samples in the accompanying table were obtained over the region 3110–3660 G. Reference to Fig. 1 will show that the spectrum of Mn^{++} is symmetrical with respect to the resonance of 1,1–diphenyl–2 picryl hydrazyl

FIG. 3. Mn^{++} spectrum in ZnS:10^{-4} Mn, Cu for 10^{-6}, 10^{-4} and $2{\cdot}5\times10^{-4}$ atomic fraction of copper.

and, therefore, the spectroscopic splitting factor, g has a value close to 2, the value of g for the hydrazyl being $2{\cdot}0036\pm0{\cdot}0002$.

When spectra of the samples containing copper (Figs. 3–5) are compared with the spectrum of manganese in cubic and hexagonal zinc sulfides (Figs. 1 and 2), it is seen that there is a change in crystal structure between $2{\cdot}5\times10^{-4}$ and $7{\cdot}5\times10^{-4}$ atomic fraction of copper. For increasing copper concentrations up to and including $2{\cdot}5\times10^{-4}$ atomic fraction, all samples have the spectrum characteristic of Mn^{++} in polycrystalline hexagonal zinc sulfide. For 5×10^{-4} Cu a large amount of cubic character is evident along with the hexagonal, and for a concentration of $7{\cdot}5\times10^{-4}$ Cu the transformation to the cubic form is practically complete.

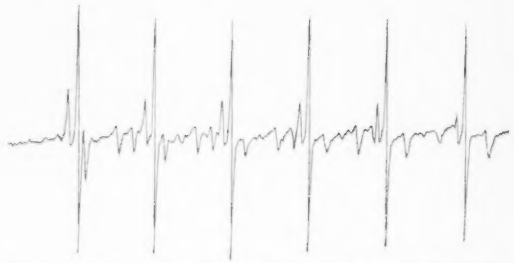


FIG. 4. Mn^{++} spectrum in ZnS:10^{-4} Mn, 5×10^{-4} Cu.

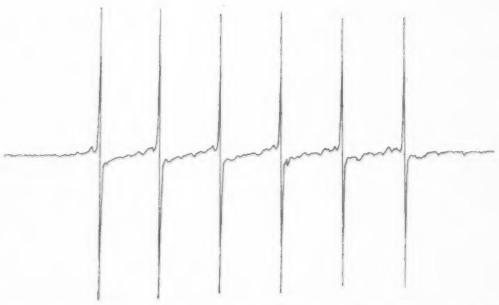


FIG. 5. Mn^{++} spectrum in ZnS:10^{-4} Mn, Cu for $7{\cdot}5\times10^{-4}$ and 10^{-3} atomic fraction of copper.

* In all figures showing EPR spectra, the abscissa represents the magnetic field increasing from left to right, and the ordinate the first derivative of absorption of microwave power.

Table 1. *ZnS*:10^{-4} *Mn, Cu,* 1 *hr in* H_2S *at* 1100°C.

EPR spectrum shown in Fig. No.	Atomic fraction of copper added	Crystal structure (EPR)	Cubic: hexagonal Ratio (X-ray)
2	none	hexagonal	—
3	10^{-6}	hexagonal	—
3	10^{-4}	hexagonal	0·011
3	$2{\cdot}5\times10^{-4}$	hexagonal	0·019
4	5×10^{-4}	hexagonal–cubic	0·059
5	$7{\cdot}5\times10^{-4}$	cubic	10
5	10^{-3}	cubic	33
1	none	cubic*	—

* Fired at 900°C to produce the cubic modification.

VOL. 13 1960

These results are summarized in Table 1 and the pertinent spectra are reproduced in Figs 3–5.

This result has been verified by X-ray diffraction analyses of five samples over the critical concentration range of copper, which also appear in Table 1. In explanation of the table, the ratio of the 3·13 to 3·31 *d* line intensities has been used as an indication of the ratio of the amount of cubic to hexagonal zinc sulfide and is referred to as the X-ray cubic:hexagonal ratio. The 3·31 reflection is due to the hexagonal modification only, but the 3·13 is common to both.

Any conclusions that might be drawn regarding the effect of copper on the phase transition in zinc sulfide might well be invalidated if manganese also affects the rate of hexagonal→cubic crystal transformation.* To check this point, three samples of the composition $ZnS{:}10^{-x}$ Mn, 5×10^{-4} Cu (where x has one of the values 5, 4 or 3) were prepared. The value of 5×10^{-4} Cu was chosen because this is a critical concentration for copper and any effects produced by the addition of manganese would be

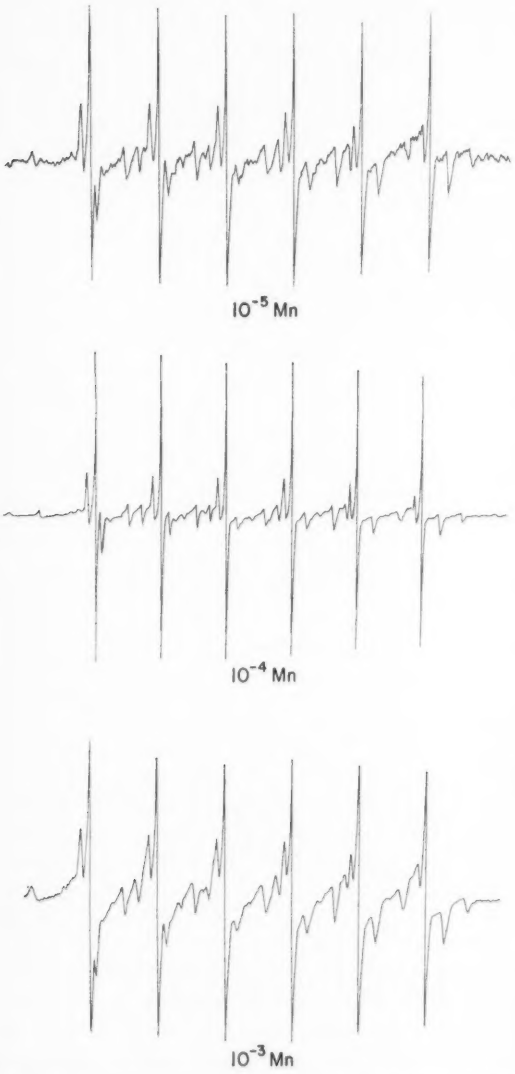


FIG. 6. EPR spectra of $ZnS{:}10^{-x}$ Mn, 5×10^{-4} Cu for $x = 3$, 4 and 5, showing negligible effect of manganese on the hexagonal–cubic phase transition.

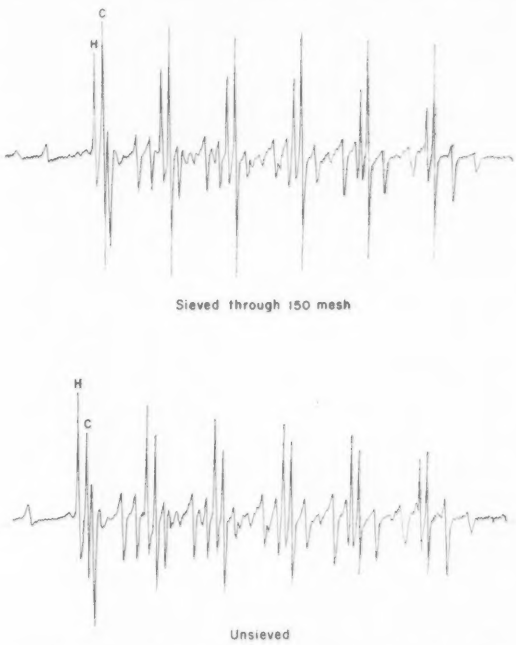


FIG. 7. Effect of sieving on phase transition in $ZnS{:}10^{-4}$ Mn fired 1 hr at 1100°C in H_2S. Peak-to-peak distances of the lines labeled C and H yield the ratio C/H, an indication of the ratio of the amount of cubic to hexagonal zinc sulfide present.

* KRÖGER[14] has shown that solid solution of relatively large amounts of MnS in ZnS lowers the hexagonal–cubic transition temperature, but for such a low concentration as 10^{-4} atomic fraction of manganese one would expect this effect to be very small.

magnified. Comparison of the three spectra (Fig. 6) show that manganese in this concentration range causes little or no change in the crystal structure of zinc sulfide. However, the effect of too large a concentration of manganese on the line width is shown in the spectrum of the sample containing 10^{-3} Mn, where spin–spin broadening is very evident.

In handling the samples it was noticed that even a relatively mild mechanical treatment, such as passing through a 150-mesh silk screen, increased the amount of the cubic phase. Fig. 7 shows the EPR spectrum of two samples of identical composition and prepared by the same technique, except that one was screened and the other was not. Close examination of the two spectra will show that the sieved specimen contains more cubic zinc sulfide than the unsieved specimen. In view of the gross similarity of the two spectra, it was felt, however, that a semiquantitative way of designating the relative amount of the cubic and the hexagonal phases was desirable. It was found that the ratio of the intensities of the two lines labelled C and H is an indication of the ratio of cubic to hexagonal content. This is purely an arbitrary designation and is not intended to give an exact ratio of the amount of cubic and hexagonal zinc sulfide present. Yet it presents a more objective way to characterize materials which are not clearly cubic or hexagonal, as all but one of the samples listed in Table 1 were. Therefore this method of designation has been adopted for all subsequent treatments of data. For the sieved sample, for example, the ratio cubic: hexagonal is 1·85, whereas the ratio for the unsieved sample is 0·88.

The EPR C/H ratios for the samples with constant copper and variable aluminum concentrations are given in Tables 2 and 3.

Table 2. *ZnS*:10^{-4} *Mn*, 10^{-3} *Cu*, *Al*, 1 *hr in* H_2S *at* 1100°C

Aluminum concentration (atomic fraction)	EPR intensity ratio C/H
0	45
$2{\cdot}5\times10^{-4}$	32
$5{\cdot}0\times10^{-4}$	14·1
$7{\cdot}5\times10^{-4}$	1·3

Table 3. *ZnS*:10^{-4} *Mn*, $7{\cdot}5\times10^{-4}$ *Cu*, *Al*, 1 *hr in* H_2S *at* 1100°C

Aluminum concentration (atomic fraction)	EPR intensity ratio C/H
0	20·2
$2{\cdot}5\times10^{-4}$	10·0
$5{\cdot}0\times10^{-4}$	3·3
$6{\cdot}0\times10^{-4}$	2·8
$6{\cdot}5\times10^{-4}$	1·9
$7{\cdot}0\times10^{-4}$	0·9

It is seen that the ratio C/H increases for decreasing aluminum concentration, indicating that the samples are becoming more cubic.

Silver added to zinc sulfide produces an effect similar to that of copper on the hexagonal→cubic phase transition. It is seen from Table 4 that the structure change occurs in the range 5×10^{-4} to 10^{-3} atomic fraction of added silver.

Table 4. *ZnS*:10^{-4} *Mn*, *Ag*, 1 *hr in* H_2S *at* 1100°C

Silver concentration (atomic fraction)	EPR intensity ratio C/H
10^{-4}	1·2
5×10^{-4}	3·3
10^{-3}	76

The crystal structure of the ZnS:9×10^{-4} Cu samples annealed with and without previous quenching is given in Table 5 below. The data were obtained by X-ray analysis. The amount of cubic crystallinity was calculated by integrating the area under the 2·70 *d* line of the cubic diffraction pattern. A standardized technique employing a General Electric XR3 diffractometer was used, all such variables as sample volume and geometry, room temperature and instrument settings being held constant. The estimated precision of the reported areas is $\pm$ 3 units, when one unit is equivalent to 1000 counts on the detector.

These data show that the highest amount of cubic crystallinity is obtained by annealing at 700°C. It is also evident that quenching with subsequent annealing produces a more complete conversion to cubic than annealing alone.

Table 5

Treatment after ½ hr at 1100°C	Amount of cubic crystallinity*		
	500°C	700°C	1000°C
Lowered to annealing temp.	53	86	53
Quenched to room temp., then annealed	73	91	65

* Integrated area under the 2·70 *d* line of X-ray diffraction pattern.

VOL. 13 1960

4. DISCUSSION

It is believed that the presented experimental results can be explained by postulating that the conversion of hexagonal zinc sulfide to cubic proceeds by the mechanism of nucleation and growth and that precipitating cuprous sulfide can supply the energy necessary for the nucleation of cubic zinc sulfide.

A phase transformation becomes thermodynamically favorable when the free energy of the new phase becomes negative relative to the free energy of the transforming phase. In the present case the free energy of the hexagonal phase is a bulk property, but the free energy of the cubic phase at the onset of transformation is a surface property. Below a critical radius the surface free energy of the embryonic cubic crystallites (nuclei) is positive relative to the free energy of the hexagonal phase, and is increasing. At the critical radius the surface free energy reaches a maximum. Above this size it begins to decrease until it is negative relative to the free energy of the hexagonal phase, at which point the conditions for growth of the cubic phase become favorable. In other words, an activation energy must be overcome before an embryo can pass into the range of stable nuclei.

The dependence of the rate of nucleation on temperature has been given by BECKER[25]:

$$N = Ce^{-Q/kT}\, e^{-A/kT}. \qquad (1)$$

Here A is the difference in free energy between the initial and final phases, i.e. it is a measure of the surface free energy of the nuclei of the new phase. Q is an activation energy related to the mobility of the atoms in the crystalline lattice. Q is not a strong function of temperature, while A increases superlinearly with temperature. The result is that below a certain temperature N increases with increasing temperature. Above this critical temperature, A overtakes the linear change in T in the denominator of $\exp[(-Q-A)/kT]$ and the nucleation rate drops.

The growth rate follows a simple exponential relation

$$G = D\exp(-B/kT), \qquad (2)$$

where neither D nor B is a strong function of temperature. Therefore, the growth rate increases approximately exponentially with temperature.

The data in Table 5 can be interpreted on the basis of a nucleation-and-growth mechanism as follows. Comparison of the amount of cubic crystallinity for the quenched and the unquenched series shows that the quenched samples show a higher degree of conversion to cubic. If nucleation at a lower temperature was not necessary for efficient phase transformation, the samples annealed without quenching should have shown greater amounts of cubic crystallinity than the quenched samples since the former spent a longer time (4 versus 3½ hr) at the high temperature where all physical and chemical changes occur more rapidly. This is clearly seen not to be the case. It is also evident that room temperature is much below the nucleation maximum, since the quenched sample annealed at 1000°C shows less cubic crystallinity than the sample quenched and then annealed at 700°C. If the optimum nucleation temperature were close to room temperature, the subsequent annealing at 1000°C would have produced a more complete transformation to the cubic form than annealing at 700°C, as the growth rate always increases with increasing temperature. As it turned out, 700°C was probably close to the nucleation maximum and yet high enough not to render the growth rate too slow. Finally, it is significant that the samples annealed at 700°C show the highest degree of conversion to the cubic modification, irrespectively whether they were previously quenched or not. Because of this and the previous point, it seems that the optimum nucleation temperature lies close to 700°C, and that at this temperature the growth is rapid enough to produce reasonably fast phase transformation.

The activation energy for nucleation is supplied in the absence of other sources by statistical fluctuations of a thermodynamic nature, in particular

configuration fluctuations. In the present system, however, it could also be supplied in part by cuprous sulfide precipitating from solid solution. Such mutual influence of two simultaneously occurring transitions or chemical reactions was first described by HEDVALL[26, 27], and is sometimes referred to as the Hedvall Effect. Thus the energy liberated by the precipitating cuprous sulfide helps to make the free energy difference, A, more negative and the rate of nucleation is increased. It appears likely that the energy evolved by the formation of a separate phase of cuprous sulfide could also go to a lowering of the diffusion activation energy, Q, with the same result, an increased nucleation rate. Once nucleation has begun, the growth of the new phase will tend to be self-sustaining, since energy is evolved in going from the hexagonal to the cubic structure. And, of course, the concentration of precipitated copper sulfide would be expected to affect the rate of nucleation, because (*a*) the more cuprous sulfide formed, the more nuclei produced and (*b*) the more copper present, the greater is the temperature range over which the sulfide precipitates. The latter effect may be of importance in assuring that cuprous sulfide will be precipitating at or near the optimum nucleation temperature for cubic zinc sulfide.

The relationship between the concentration of copper, silver and aluminum and the rate of the hexagonal–cubic phase transformation can be understood on the basis of the following considerations. As it is quite generally the case that the solubility of a solute in a solvent increases with temperature, it is reasonable to assume that most of the small amount (e.g. 10^{-3} atomic fraction) of added copper will be dissolved in zinc sulfide at the firing temperature of 1100°C. On cooling, some of the copper will precipitate from solution as cuprous sulfide. The quantity retained in solution will correspond, as a first approximation,* to its room-temperature solubility. It is believed that in the absence of chemical coactivating species the solution of copper in zinc sulfide occurs by the process of self-coactivation.[28, 19] This view is based on the observation that, although copper acts as an acceptor in zinc sulfide, the introduction of copper alone does not produce an uncompensated p-type material under normal preparative conditions. Instead physical coactivating species, most likely sulfur vacancies, are formed, which act as donors and compensate the introduced copper acceptors. The amount of copper which can be retained by this self-coactivation mechanism is limited to less than about 2×10^{-4} atomic fraction. Evidently, at higher copper concentrations, the lattice strain due to the introduced vacancies becomes too large and the further introduction of copper becomes energetically unfavorable. The excess copper under such conditions is ejected from solid solution as the sample is cooled from the firing temperature to room temperature. The precipitated copper sulfide lends the zinc sulfide a grayish appearance which can be partly removed by washing with sodium cyanide. If aluminum is introduced simultaneously with copper, copper sulfide does not precipitate even at concentrations as high at 10^{-2} atom fraction.[29] The samples do not appear gray, but acquire a yellowish appearance. This effect of aluminum has been interpreted on the basis of its acting as a compensating donor for the copper acceptors. The atomic radii of copper and aluminum are not too different from that of zinc, so the lattice strain introduced by substitution of these atoms for zinc is not excessive. Therefore the increase in entropy resulting from the formation of solid solution of copper and aluminum in zinc sulfide, and the lowering of internal energy due to the transfer of electrons from aluminum donors to copper acceptors keeps the introduction of these species into zinc sulfide energetically favorable to a relatively high concentration.

According to the data of Table I, 5×10^{-4} atomic fraction is the concentration of copper at which rate of transformation to cubic becomes noticeable. As aluminum is able to increase the solubility of copper, it would be expected that, in samples containing aluminum, the conversion to cubic would not become appreciable until the concentration of copper in excess of that of aluminum is about 5×10^{-4} atomic fraction. That this is indeed the case is evident from Tables 2 and 3. Using the C/H ratio of approximately ten as a measure of a starting phase transformation it is seen that with 10^{-3} atomic fraction copper the transformation is appreciable at 5×10^{-4} atomic

* Actually the quantity of copper retained in solution at room temperature will correspond to the solubility at a higher temperature at which the precipitation rate becomes negligible.

VOL. 13 1960

fraction aluminum, and that with $7{\cdot}5\times10^{-4}$ atomic fraction copper it occurs at $2{\cdot}5\times10^{-4}$ atomic fraction aluminum. Thus, in both cases the onset of rapid transformation occurs at an excess copper concentration of 5×10^{-4} atomic fraction.*

Considering further, that the solubility (or, more precisely, solubility at the temperature where no more perceptible atomic rearrangements can take place) of self-coactivated copper in zinc sulfide is about 2×10^{-4} atomic fraction, it means that the precipitation of only a small amount of copper is required to initiate a rapid phase transformation of hexagonal zinc sulfide to cubic. In the case of silver, the situation is similar. The figures in Table 4 show that the onset of rapid phase transformation occurs between 5×10^{-4} and 10^{-3} atomic fraction silver. As the solubility of self-coactivated silver in zinc sulfide is not known, it is hard to estimate the actual amount of silver sulfide precipitate present under such conditions. Qualitatively, however, it appears that, within a factor of two, the amount of silver sulfide precipitate which triggers the transformation is equal to that of cuprous sulfide.

If there is a relation between the crystal structure of cuprous sulfide or silver sulfide and the ability of each of these compounds to catalyze the formation of cubic zinc sulfide (beyond the unspecific thermodynamic effects described earlier), its function is not at all clear. The high-temperature form of Cu_2S (chalcocite) is hexagonal[30, 31] with space group D^4_{6H} ($C6/mmc$). There is another high-temperature phase, however, of the composition Cu_9S_5, which is called digenite and was for a long time thought to possess a face-centred cubic structure.[32, 33] Recently DONNAY[34] has shown that digenite is actually rhombohedral, pseudocubic. To further confuse things, the compound cupric sulfide will coexist with Cu_9S_5 if the material is prepared in but a slight excess of sulfur.[34, 35] In the other direction of a slight excess of copper, a mixture of Cu_2S and Cu_9S_5 is produced. Inasmuch as the composition of the copper sulfide phase associated with zinc sulfide is not known and the amount present is below the limit of detection by X-ray diffraction, one cannot even hazard a guess at its structure. In the case of silver sulfide the situation is also confusing. There is a modification which is stable between about 180 and 625°C, which is reported[32] to have a body-centered cubic structure. Above 625°C Ag_2S II, the structure of which has not been determined, exists.[36] Both cuprous sulfide and silver sulfide have low-temperature modifications: below 105°C cuprous sulfide is orthorhombic, pseudo-hexagonal[31] and below 170°C silver sulfide is monoclinic, pseudo-rhombic.[37]

It is of interest to note that sulfides of divalent elements of suitable size, which can be incorporated into zinc sulfide without coactivation and which, therefore, are more soluble in zinc sulfide than the sulfides of monovalent elements, have a completely different effect on the transformation of zinc sulfide. ALLEN and CRENSHAW[13] have shown that with increasing concentration of iron in zinc sulfide, the hexagonal–cubic transition is moved to lower temperatures. Here, evidently, a true change in the transformation point is encountered, while in the case of excess copper and silver, the transformation point itself is not affected merely the rate of transformation. This does not mean that the part of copper or silver staying dissolved in zinc sulfide does not influence the transformation point. It could well be that dissolved copper or silver actually tends to lower the hexagonal–cubic transformation point, as was found to be the case with iron; but it is expected that the effect would be slight, because the concentration of solute is so low.

5. SUMMARY

The formation of a separate phase of copper sulfide or silver sulfide in polycrystalline ZnS:Cu, ZnS:Ag and ZnS:Cu, Al triggers the rapid transformation of hexagonal to cubic zinc sulfide. It is believed that after firing cuprous sulfide precipitates

* It could be argued that copper sulfide did not influence the transformation rate of hexagonal zinc sulfide to cubic, but influenced the growth of the zinc sulfide crystallites in such a way as to produce the experimentally found proportions of cubic and hexagonal material at the firing temperature. Aside from the self-consistency of the above presented interpretation, this is not likely for two reasons. First, if anything, copper sulfide would be expected to lower the hexagonal-cubic transition temperature (see below) and therefore certainly not to inhibit the formation of hexagonal structure at 1100°C. Second, experiments have shown that the extent of conversion to cubic depends on the time of annealing at temperatures below the hexagonal-cubic transition point, which means that transformation is indeed proceeding at these temperatures.

from solution in hexagonal zinc sulfide and in so doing supplies the energy necessary for nucleation of cubic zinc sulfide. Once nucleation has begun, growth of the cubic phase will be self-sustaining, because energy is released in going from the hexagonal to the cubic structure. Supporting data for the postulated mechanism of nucleation and growth is provided by X-ray diffraction data on the effect of quenching and annealing on the crystal structure of ZnS:9×10^{-4} Cu phosphors. It was found that the optimum nucleation temperature lies close to 700°C and that at this temperature the growth of the cubic phase is reasonably fast.

Acknowledgement—The authors wish to gratefully acknowledge the contribution of Miss J. R. Cooper in performing the X-ray diffraction analysis.

REFERENCES

1. Allen E. T. and Crenshaw J. L., *Z. anorg. Chem.* **79**, 130 (1913).
2. Frondel C. and Palachi C. *Amer. Min.* **35**, 29 (1950).
3. Strock L. W. and Brophy V. A. *Amer. Min.* **40**, 94 (1955).
4. Buck D. C. and Strock L. W., *Amer. Min.* **40**, 192 (1955).
5. Schleede A. and Gantzchow H. *Z. phys. Chem.* **106**, 37 (1923).
6. Fuller M. L., *Phil. Mag.* **8**, 658 (1929).
7. Bridgman P. W. *Amer. J. Sci.* **237**, 7 (1939).
8. Frey F. *Ann. Phys. Lpz.* **2**, 147 (1948).
9. Kullerud G., *Norsk Geol. Tideskr.* **32**, 61 (1953).
10. Strock L. W., *Sylvania Techn.* **8**, 71 (1955).
11. van Wieringen J. S., *Physica, s'Grav.* **19**, 397 (1953).
12. Smith F. G. and Hill V. G., *Acta Cryst.* **9**, 821 (1956).
13. Allen E. T. and Crenshaw J. L., *Amer. J. Sci.* **34**, 348 (1912).
14. Kröger F. A., *Z. Krist*, **100**, 543 (1939); **102**, 132 (1940).
15. Hershberger W. D. and Leifer H. N., *Phys. Rev.* **88**, 714 (1952).
16. Schneider E. E. and England T. S., *Physica s'Grav.* **17**, 221 (1951).
17. Matarrese L. M. and Kikuchi C., *J. Phys. Chem. Solids* **1**, 117 (1956).
18. Keller S. P., Gelles I. L. and Smith W. V., *Phys. Rev.* **110**, 850 (1958).
19. Aven M. and Potter R. M., *J. Electrochem. Soc.* **105**, 134 (1958).
20. Bleaney B. and Stevens K. W. H., *Rep. Progr. Phys.* **16**, 108 (1953).
21. Bowers K. D. and Owen J., *Rep. Progr. Phys.* **18**, 304 (1955).
22. Bagguley D. M. S. and Owen J., *Rep. Progr. Phys.* **20**, 304 (1957).
23. Ingram D. J. E., *Spectroscopy at Radio and Microwave Frequencies*. Butterworths, London (1955) and *Free Radicals as Studied by Electron Spin Resonance*. Academic Press, New York (1958).
24. Gordy W., Smith W. V. and Trambarulo R. F., *Microwave Spectroscopy*. John Wiley, New York (1953).
25. Becker R., *Ann. Phys. Lpz.* **32**, 128 (1938).
26. Hedvall T. A. and Sjoman D., *Z. Elektrochem.* **37**, 130 (1931).
27. Hedvall T. A. and Olson K., *Z. anorg. Chem.* **243**, 237 (1940).
28. Bowers R. and Melamed N. T., *Phys. Rev.* **99**, 1781 (1955).
29. Froelich H. C., *J. Elektrochem. Soc.* **100**, 496 (1953).
30. Buerger M. J. and Buerger N. W., *Amer. Min.* **29**, 55 (1944).
31. Belov N. V. and Butuzov V. P., *C. R. Acad. Sci. U.S.S.R.* **54**, 717 (1946).
32. Rahlfs P., *Z. phys. Chem.* B**31**, 157 (1936).
33. Buerger N. W., *Econ. Geol.* **36**, 17 (1941); *Amer. Min.* **27**, 712 (1942).
34. Donnay G., Donnay J. D. H. and Kullerud G., *Amer. Min.* **43**, 228 (1958).
35. Hansen M. and Anderko K., *Constitution of Binary Alloys*, p. 620. McGraw-Hill, New York (1958).
36. Hansen M. and Anderko K., *Constitution of Binary Alloys*, p. 46. McGraw-Hill, New York (1958).
37. Ramsdell L. S., *Amer. Min.* **28**, 401 (1943).

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 13. pp. 65–70. Printed in Great Britain.

ANISOTROPIC DIFFUSION OF COPPER INTO BISMUTH TELLURIDE

R. O. CARLSON

General Electric Research Laboratory, Schenectady, New York

(*Received* 3 *August* 1959)

VOL. 13 1960

Abstract—Bismuth telluride (Bi_2Te_3) has a layer structure (. . . – – Te – Bi – Te – Bi – Te – – Te – Bi – Te – Bi – Te – – Te . . .) and readily cleaves perpendicular to the *c*-axis of the rhombohedral crystal. A marked difference has been found between the penetration of radioactive copper parallel to and perpendicular to the cleavage planes in Bi_2Te_3. $D_{\parallel} = D_0 \exp -(E/kT) = 0{\cdot}0034 \exp(-0{\cdot}21\ \text{eV}/kT)$, $D_{\perp} = 0{\cdot}071 \exp(-0{\cdot}80\ \text{eV}/kT)$. At room temperature, $D_{\parallel} \sim 10^{-6}$ cm^2/sec, while extrapolated $D_{\perp}$ is more than eight decades lower at $\sim 3 \times 10^{-15}$ cm^2/sec. Minute cracks, except perhaps on an atomic scale, are not believed responsible for the fast parallel diffusion. The fast parallel diffusion can be rationalized on the basis that in this direction, between adjacent tellurium layers the copper moves through a region of relatively weak electrostatic bonding forces and large layer spacing. In any other direction, covalent and ionic bonding between tellurium and bismuth atoms would make penetration more difficult. Zener's theory is shown to give a reasonable fit to the D_0 values, requiring a negligible entropy term for parallel diffusion and a small entropy term, approximately the Boltzmann factor k, for perpendicular diffusion.

BISMUTH telluride (Bi_2Te_3) has attracted attention in recent years as a thermoelectric material. Since stoichiometric bismuth telluride is *p*-type, a number of elements were tried as additives in order to convert it to an *n*-type semiconductor. Copper proved to be a suitable donor impurity. However, on standing, even at room temperature, copper diffuses to the surface of the bismuth telluride and possibly also to internal precipitates and the material becomes less *n*-type. Copper also was unintentionally added if the solder used in making junctions to bismuth telluride was not relatively free of copper.[1, 2]

The structure of bismuth telluride is well established[3] and described as rhombohedral with space group $D_{3d}^5 (R\bar{3}m)$. The atoms are arranged in layers in the order . . . – – Te – Bi – Te – Bi – Te – – Te . . . in the rhombohedral (111) direction. The marked anisotropy in crystal structure is reflected in measurements of thermoelectric power, electrical resistivity, thermal conductivity and Hall coefficient,[4] and easy cleavage into mica-like sheets.

Polycrystals were grown from a bismuth telluride melt (slight excess of tellurium) in a quartz crucible, adapting the melted layer technique[5] to make the crystals as uniform as possible. The volume above the melt was maintained near 500°C which is about 100 degrees below the melt temperature. A quartz ring seal permitted rotation and pulling of the growing crystal, at the same time serving as a high impedance leak for the tellurium vapor. The polycrystals consisted of a few crystallites oriented along the growth axis. In some crystals, there were regions in which the usual ellipsoidal shape was flattened out. From these regions, it was possible to cleave parallel-sided samples with dimensions about one to two cm^2 area and several mm thickness. The surfaces of these samples were parallel to the growth axis, perpendicular to the *c*-axis, and correspond to (0001) planes when referred to the hexagonal axis. These samples were labeled as "perpendicular" samples since the copper was diffused in from the surface perpendicular to the cleaved face.

A carborundum saw was used to make cuts transverse to the growth direction of crystals with ellipsoidal cross section. For the cutting operation, the crystal was imbedded in a plastic to prevent cleaving or cracking by the saw motion. These

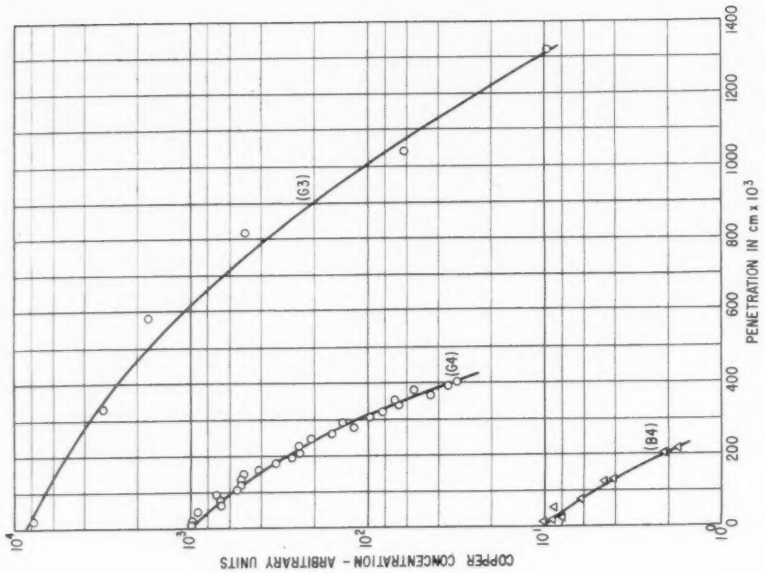


FIG. 2. Copper penetration into Bi_2Te_3 in the "parallel" direction, i.e. parallel to the cleavage plane. The curves are conjugate error functions with diffusion temperatures and times and "best-fit" diffusion coefficients as listed in Table 1. The data points and curves for these three typical samples have been adjusted vertically to avoid overlapping.

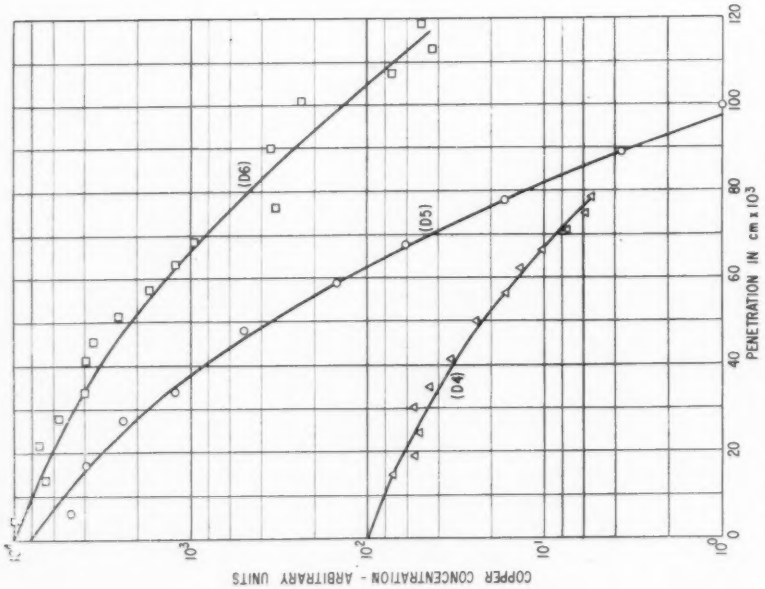


FIG. 1. Copper penetration into Bi_2Te_3 in the "perpendicular" direction, i.e. perpendicular to the cleavage plane or along the *c*-axis of the crystal. The curves are conjugate error functions with diffusion temperatures and times and "best-fit" diffusion coefficients as listed in Table 1. The data points and curves for these three typical samples have been adjusted vertically to avoid overlapping.

samples were called "parallel" samples since diffusion in from the cut end surfaces was along the cleavage direction.

To follow the copper penetration, Cu-64 ($\beta\pm$, 12·8 hr) was electrolytically deposited from copper nitrate solution on a cleaved face of a perpendicular sample or a cut face of a parallel sample. Samples

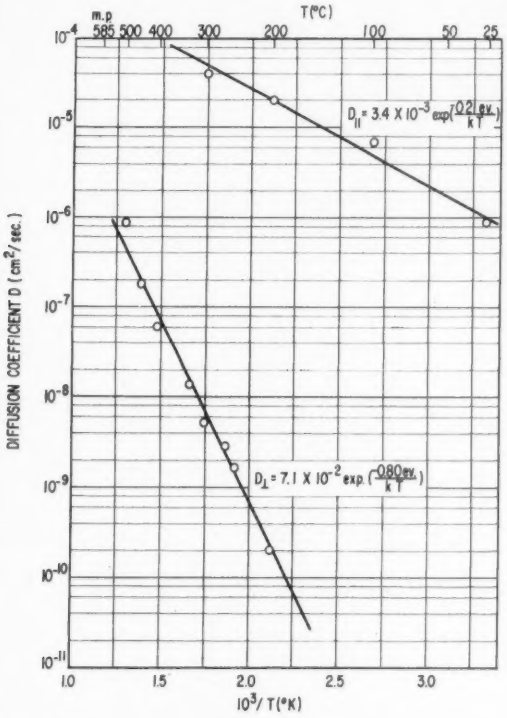


FIG. 3. Diffusion coefficient vs. reciprocal temperature of copper in Bi_2Te_3 for two directions of penetration, parallel to the cleavage plane or perpendicular to the cleavage plane.

were placed in furnaces under a hydrogen atmosphere. On removal, the samples were mounted and grindings made on emery paper in a small precision grinder. The grindings were counted in a plastic scintillation counter to establish the dependence of copper concentration on depth below the surface of the sample.

The penetration data for the perpendicular samples, as shown in Fig. 1, were fitted by the conjugate error functions (solid lines), in some cases over three to four decades. Penetration data for the parallel samples showed greater scatter of points, the curves in Fig. 2 being those for the better samples. This scatter was largely due to flaking and chipping of the parallel samples during grinding even though they were supported by spring clips on the grinding head. In the parallel direction, the diffusion is rapid at room temperature, increasing slowly with temperature. The diffusion coefficients calculated for both directions of penetration are tabulated in Table 1 and plotted in Fig. 3 as a function of reciprocal temperature. Surface solubilities were in the range 10^{18}–10^{19} atoms of copper/cm³, but were not accurately determined because only the central portions of the cleaved or cut faces were copper coated in order to avoid edge effects. Though a few samples were allowed to diffuse long enough to saturate them, their bulk copper concentrations were again low because not enough copper was deposited at the surfaces for bulk saturation.

Although the transverse cutting out of the

Table 1. *Diffusion temperatures and times and computed diffusion coefficients*

No.	Temperature (°C)	Time	Diffusion coefficient (cm^2/sec)
(A) Diffusion parallel to cleavage planes			
*F*2	25	29 hr	$\sim 9 \times 10^{-7}$
*G*3	25	2 hr	$9{\cdot}0 \times 10^{-7}$
*F*1	25	2 hr	$\sim 6 \times 10^{-7}$
*G*2	100	15 min	$\sim 7 \times 10^{-6}$
*G*4	200	15 min	$2{\cdot}0 \times 10^{-5}$
*B*1	285	5 min	$\sim 5 \times 10^{-5}$
*B*4	300	5 min	4×10^{-5}
*E*3b	450	60 min	$(\sim 8 \times 10^{-5})$
*E*5b	450	32 min	$(\sim 1{\cdot}1 \times 10^{-4})$
(B) *Diffusion perpendicular to cleavage planes*			
*D*2	200	72 hr	$\sim 2 \times 10^{-10}$
*E*3	250	64 hr	$1{\cdot}6 \times 10^{-9}$
*D*3	263	70 hr	$2{\cdot}8 \times 10^{-9}$
*D*5	300	17 hr	$5{\cdot}2 \times 10^{-9}$
*D*6	330	18 hr	$1{\cdot}3 \times 10^{-8}$
*D*4	400	203 min	$6{\cdot}0 \times 10^{-8}$
*C*2	450	200 min	$1{\cdot}8 \times 10^{-7}$
*E*3a	450	60 min	$1{\cdot}8 \times 10^{-7}$
*E*5a	450	32 min	$\sim 1{\cdot}0 \times 10^{-7}$
*D*1	500	92 min	$8{\cdot}9 \times 10^{-7}$

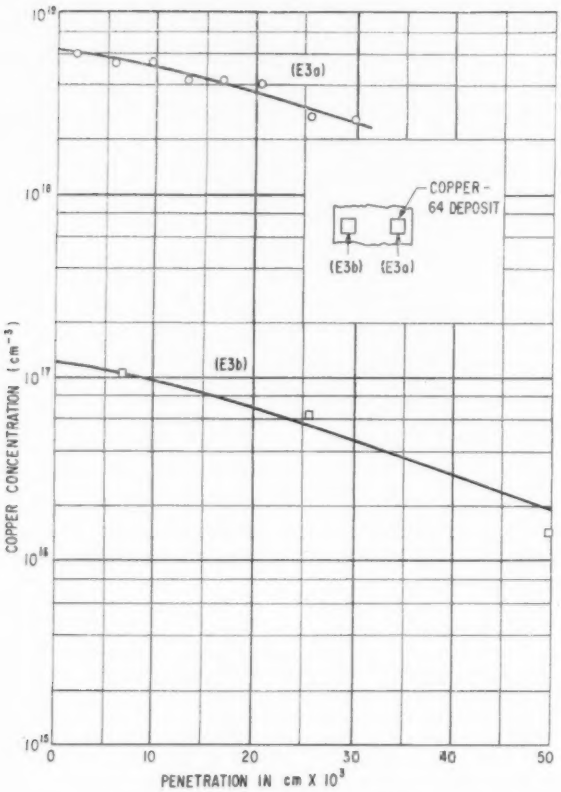


FIG. 4. Copper penetration into Bi_2Te_3 in the "perpendicular" direction at 450°C for 1 hr. Sample *E*3a was under the copper-64 coated area, *E*3b approximately 1·8 cm away on the cleavage face. Curves are drawn for $D_\perp = 1{\cdot}8 \times 10^{-7}$ cm²/sec. The calculated diffusion along the surface is $\sim 8 \times 10^{-5}$ cm²/sec.

parallel samples was carefully done, and no cracks could be observed visually away from the edges, the possibility that the fast diffusion is due to migration along minute cracks could not be completely discounted. Etching for a long period in an HCl–HNO_3 mixture[1] did bring out a few apparent cracks in some samples but this may be due to preferential etching in certain crystallographic directions. To test whether cracks were already built into the crystals and not just introduced by subsequent handling, very short crystals were grown and the copper-64 applied directly to one end of these crystals. The curve for room temperature in Fig. 2 (sample *G*3) represents data on a short crystal which fits very well with the other points taken on sawed out sections of large crystals.

As a further check on the question of diffusion along cracks, a determination was made of the diffusion rate along the surface of a cleaved face. Copper was applied on only a small part of the cleaved face of a perpendicular sample (No. *E*3), the rest of the face being masked off. Then the sample was held at 450°C for an hour. On removal, two regions were cut out for grinding, one immediately under the copper-64 coated region, the other 1·6 to 2·0 cm away. Penetration data are plotted in Fig. 4 and for both regions $D = 1{\cdot}8 \times 10^{-7}$

VOL. 13 1960

cm^2/sec. The ratio of surface concentrations is 52, and taking the average separation of regions as 1·8 cm, one calculates the surface diffusion coefficient as $\sim 8 \times 10^{-5}$ cm^2/sec. A second sample (No. *E*5) is listed in Table 1 as giving $\sim 11 \times 10^{-5}$ cm^2/sec. Within experimental error, these values fit the data for the parallel diffusion coefficient extrapolated to 450°C. Thus if cracks act like the surface does, and this seems a reasonable assumption since the cracks should be along the direction of easy cleavage, then diffusion along cracks must be like the measured parallel diffusion. But if a relatively few large cracks provide a short circuiting mechanism for diffusion, there would presumably be a wide scattering of our data since the density of cracks would vary from sample to sample. Since the data are not widely scattered, then any cracks must be on a near atomic scale and inherent in the nature of the bismuth telluride itself.

In the layer structure of Bi_2Te_3 (. . . Te_{II} – – Te_{II} – Bi – Te_I – Bi – Te_{II} – – Te_{II} – Bi – Te_I . . .), the various tellurium layers must be distinguished because their nearest neighbor configurations are different. Within each bismuth layer, the atoms of bismuth are 4·38 Å apart while within each tellurium layer, the tellurium atoms are 4·38 Å apart. The spacings between atoms in different layers are Te_{II} – Te_{II} as 3·72 Å, Te_{II} – Bi as 3·04 Å, and Bi – Te_I as 3·24 Å[6] The shortest spacings between the layers themselves are as follows:

Te_{II} – Te_{II}, 2·63 Å; Te_{II} – Bi, 1·70 Å; Bi – Te_I, 2·03 Å.

For the parallel diffusion case, the large diffusion coefficients suggest interstitial diffusion. The copper may move through the lattice as an ion, for which the Pauling radius is given as 0·96 Å, or, less likely, as an atom, with radius of 1·27Å. If one considers the nature of the bonds in Bi_2Te_3, that between neighboring bismuth and tellurium atoms appears to be covalent and ionic, whereas that between Te_{II} atoms in adjacent layers is likely to be by the relatively weak van der Waals forces,[6,7] with perhaps some weak covalent bonding.[8] In the layers themselves, the bonding is also weak. Thus in the parallel direction the relatively larger spacing and weaker electrostatic forces between the Te_{II} layers indicates that this is the region through which the copper could diffuse rapidly.

Various expressions have been derived for calculating D_0[9,10] for comparison with experimental values of D_0. In ZENER's theory[11] D_0 can be written as $\lambda a_0^2 \nu \exp(\Delta S/k)$. Here λ is a geometrical and probability factor which takes account of and weights the several paths the diffusing particle can take, a_0 is the lattice constant, ν the vibration frequency of the diffusing particle in the potential field between interstitial sites and approximately equal to $\sqrt{(E/2ma^2)}$—E the activation energy in the expression $D = D_0 \exp(E/kT)$, m the copper mass, and a the jump distance—ΔS an entropy term to account for the change in the lattice strain energy due to the diffusing particle, and k the Boltzmann factor.

Considering the parallel diffusion case first, an interstitial site between the Te_{II} layers has six symmetrical sites available to it at a jump distance of a_0. Weighting the jump distances appropriately in a forward direction, one calculates λ as 3/2.

The lattice constant for Bi_2Te_3 is 4·38 Å. $\nu = 0{\cdot}91 \times 10^{12}$ sec^{-1}. Then the calculated $D_0 = 2{\cdot}6 \times 10^{-3} \exp(\Delta S/k)$ compared to the experimental value of $3{\cdot}4 \times 10^{-3}$, with an estimated uncertainty in fitting the Fig. 3 data of $\pm\ 1 \times 10^{-3}$. Thus the entropy term ΔS is essentially zero or at most about $\frac{1}{2}k$, a result that is reasonable considering the open structure of Bi_2Te_3 and the relatively small distortion in the lattice by copper, especially if present as an ion.

Considering now the perpendicular diffusion case, we note that when viewed at an angle to the *c*-axis, bismuth telluride has almost a cubic structure.[12] Since the centers of these cubes offer the biggest interstitial sites, "perpendicular" diffusion will most likely take place along the axes of these cubes. From a given interstitial site six directions are available, three with a component downward along the *c*-axis, three with a component upward along the *c*-axis. The jumps from one "cubic" cell to the next are not equally spaced, nor necessarily equally difficult. Five successive jumps are required to go from between one pair of Te_{II} layers to the next pair. Two of these jumps involve passing two Bi–Te bonds, the other three passing only one Bi–Te bond. To simplify the calculation of λ, it was assumed that the diffusion was limited by the former two jumps for which the jump distance is 3·14 Å and the component along the *c*-axis is 1·87 Å. The factors that enter into D_0 are as follows: $\lambda a_0^2 = 0{\cdot}55 a_0^2 \times (10{\cdot}09/3{\cdot}73)^2$, the

second factor accounting for the three easy jumps and two difficult jumps in each sequence; $\nu = 2{\cdot}48 \times 10^{12}$. Then the calculated $D_0 = 1{\cdot}9 \times 10^{-2} \exp(\Delta S/k)$ cm^2/sec, compared with the experimental value of $7{\cdot}1^{+10}_{-3} \times 10^{-2}$. ΔS is then $1{\cdot}3^{+0{\cdot}9}_{-0{\cdot}5}k$ or $1{\cdot}1^{+0{\cdot}8}_{-0{\cdot}4} \times 10^{-4}$ev/deg. ZENER[10] relates ΔS to the temperature dependence of the elastic moduli, which unfortunately are not known for Bi_2Te_3. However, using ZENER's expression for ΔS as $\sim\beta E/T_m$ where E is the activation energy (0·80 eV), T_m the melting point (858°K) and β involves the elastic moduli as $-d(\mu/\mu_0)/d(T/T_m)$ and solving for β, we find for Bi_2Te_3 that $\beta \sim 0{\cdot}12^{+0{\cdot}08}_{-0{\cdot}05}$. Comparing with Table 1 in ZENER's paper, these β values are low as compared with most of the metals listed, but in the range of more open crystal structures like Ge ($\beta = 0{\cdot}13$) and Si (0·06).[13]

Thus ZENER's theory reasonably well fits the experimental data, with the entropy factor ΔS approximately zero for diffusion along the cleavage plane and approximately k for diffusion perpendicular to the cleavage plane.

Few cases of diffusion anisotropy[9] have been reported in the literature and these largely involve self diffusion in metals such as tin, zinc and bismuth.* For these cases, ZENER's theory has not given a good fit to the D_0 values and, in fact, would require negative entropy factors. It is interesting to note that no diffusion anisotropy was found for copper penetration into hexagonal type cadmium sulfide.[14]

* The element tellurium also has a very anisotropic crystal structure, but some exploratory measurements of copper-64 diffusion into single crystal tellurium by A. G. TWEET (private communication) suggest that the diffusion may be rapid without, however, any marked dependence on orientation.

Acknowledgements—I should like to express my appreciation to F. H. HORN, who developed the crystal growing techniques, to R. F. REIHL, who grew the Bi_2Te_3 crystals, to G. BROWER, who assisted in cutting, cleaving and grinding the crystals, and to A. G. TWEET and E. M. PELL who greatly aided in the interpretation and understanding of the data.

REFERENCES

1. DRABBLE J. R., GROVES R. D. and WOLFE R., *Proc. Phys. Soc.* B**71**, 430 (1958).
2. GOLDSMID H., J. *Proc. Phys. Soc.* B**71**, 633 (1958).
3. LANGE P. W., *Naturwissenschaften* **27**, 133 (1939). FRANCOMBE M. H., *Brit. J. Appl. Phys.* **9**, 415 (1958).
4. GOLDSMID H. J., SHEARD A. R. and WRIGHT D. A., *Brit. J. Appl. Phys.* **9**, 365 (1958). DRABBLE J. R., *Proc. Phys. Soc.* B**72**, 380 (1958).
5. HORN F. H., *Trans. Electrochem. Soc.* **105**, 393 (1958).
6. AIRAPEYANTS S. V. and EFIMOVA B. A., *J. Tech. Phys., Moscow* **28**, 1768 (1958).
7. DRABBLE J. R. and GOODMAN C. H. L., *J. Phys. Chem. Solids* **5**, 142 (1958).
8. MOOSER E. and PEARSON W. B., *J. Phys. Chem. Solids* **7**, 65 (1958).
9. BARRER R. M., *Diffusion in and through Solids*, pp. 268–305. Cambridge University Press (1941). SEEGER A., *Handbuck der Physik* Bd. 7, Teil 1, S. 407–426. Springer, Berlin (1955).
10. ZENER C. in *Imperfections in Nearly Perfect Crystals*, pp. 289–314. Wiley, New York (1951).
11. WERT C. and ZENER C., *Phys. Rev.* **76**, 1169 (1949). (Their notation is not the same as in this article.)
12. WRIGHT D. A. in *Semiconductors and Phosphors*, pp. 477–481. Interscience Publ., New York (1958). Three figures in this paper are photographs of a model of Bi_2Te_3 and show the layer structure when viewed perpendicular to the c-axis, the hexagonal structure when viewed along the c-axis direction, and the quasi-cubic structure at an angle to the c-axis.
13. PELL E. M., private communication.
14. CLARKE R. L., *J. Appl. Phys.* **30**, 957 (1959).

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 13. pp. 71–80. Printed in Great Britain.

MECHANISM OF ANTIFERROMAGNETISM IN DILUTE ALLOYS

A. W. OVERHAUSER

Scientific Laboratory, Ford Motor Company, Dearborn, Michigan

(*Received* 2 *October* 1959)

VOL. 13 1960

Abstract—A mechanism for the antiferromagnetic ordering of a dilute paramagnetic solute in a metal is proposed and discussed in relation to the phenomena that occur in copper–manganese alloys. Long-range antiferromagnetic order results from a static spin-density wave in the electron gas of the metal. This new state of the gas is dynamically self-sustaining as a result of the exchange potentials arising from the spin-density distribution. The paramagnetic solute atoms are then oriented by their exchange interaction with the spin-density wave. The resulting interaction energy more than compensates the increase in energy associated with the formation of the spin-density wave. The theory predicts correctly the magnitude and concentration-dependence of the critical temperature, the anomalous low-temperature specific heat and the anomalous electrical and magnetic properties of the alloys.

1. INTRODUCTION

RECENT experimental studies[1–4] of dilute alloys of manganese in copper have revealed several interesting and unusual properties. Among these is a low-temperature maximum in the magnetic susceptibility—behavior that is characteristic of an antiferromagnetic transition. On the other hand, the temperature-dependence of the reciprocal susceptibility at higher temperatures has a positive intercept on the temperature axis (when extrapolated)—behavior that is characteristic of a ferromagnetic transition. Low-temperature specific-heat measurements[5] on the alloys show an excess specific heat (compared to pure copper) in the temperature region of the susceptibility maximum. Very likely a coöperative phenomenon of some sort is occurring. The transition region is broad, however, presumably as a result of inhomogeneities associated with a random distribution of the paramagnetic solute. The transition temperature is approximately proportional to the solute concentration, being $\sim 25°K$ for 1 atomic per cent manganese. Small magnetic remanences, corresponding to ~ 1 per cent magnetic alignment of the manganese, can be induced. It is difficult to decide experimentally whether the low-temperature order is basically antiferromagnetic or whether for example, it consists of small ferromagnetic domains coupled antiferromagnetically.[6] Both the entropy associated with the excess specific heat and the susceptibility above the transition temperature indicate that spin state of the manganese is $S = 2$.

Finding a mechanism of long-range order in a dilute alloy is not easy at first sight. SATO *et al.*[7] have recently pointed out that long-range order can occur only if the concentration of magnetic solute is greater than 10 or 15 atomic per cent, assuming only short-range interactions are operative. They have emphasized that the familiar molecular-field approximation is completely erroneous when the product of coördination number and solute concentration is not larger than unity. Imagining a superlattice structure associated with a randomly distributed species is an even more formidable task. It seems essential to postulate a long-range interaction, or its equivalent, for an approach to the problem in copper–manganese alloys.

HART[8] has shown that interactions between the spins of solute atoms in a metal are always screened by the dynamical readjustment of the conduction electrons. YOSIDA[9] has exhibited this behavior in detail by means of perturbation theory. Although

the resulting interaction is not restricted to nearest neighbors, the long-range part of it oscillates in sign and diminishes rapidly in magnitude with distance. Consequently this pseudo-exchange interaction seems to have meager potentiality for producing coöperative phenomena in a dilute alloy.

Lack of an alternative suggests that the phenomena of interest result from the exchange interaction between the spins of the paramagnetic solute and those of the conduction electrons. The exchange Hamiltonian can be approximated:

$$H' = -(G/N)\sum_{i,j} \boldsymbol{s}_i \cdot \boldsymbol{S}_j \delta(\boldsymbol{r}_i - \boldsymbol{R}_j), \qquad (1)$$

where i labels the spin and position operators of the conduction electrons and j labels the spin and lattice site of the paramagnetic solute atoms. Employment of the delta function in equation (1) is equivalent to neglecting the dependence of the exchange integrals on the wave vectors of the conduction electrons. The magnitude of the coefficient G can be estimated from the spectroscopic-term values associated with the s–d exchange interaction of the $(3d)^4 4s$ or $(3d)^6 4s$ manganese configurations, which yield a value

$$G \sim 0{\cdot}4 \text{ eV}. \qquad (2)$$

The number of atoms per unit volume is N. The Hamiltonian, (1), has been used by YOSIDA in deriving the effective interaction[9] between pairs of manganese spins, and in calculating the contribution of spin-disorder scattering to the resistivity of the alloys.[10] In the present work the interaction, (1), will be used only as the origin of an effective field for orienting the spins of the solute atoms. As will be shown in the following sections, long-range antiferromagnetic order is caused by a static spin-density wave in the conduction-electron gas:

$$\boldsymbol{s}(\boldsymbol{R}) = bN\boldsymbol{\epsilon} \cos \boldsymbol{q} \cdot \boldsymbol{R}, \qquad (3)$$

where b is a (dimensionless) amplitude, $\boldsymbol{\epsilon}$ is a unit polarization vector, and $\boldsymbol{s}(\boldsymbol{R})$ is the electron spin density (in units of $\hbar$), given by:

$$\boldsymbol{s}(\boldsymbol{R}) = \langle \sum_i \boldsymbol{s}_i \delta(\boldsymbol{r}_i - \boldsymbol{R}) \rangle, \qquad (4)$$

the expectation value of the total spin-density operator. The effective field $\boldsymbol{H}_j$ acting on the solute spin $\boldsymbol{S}_j$, as defined by the interaction energy,

$$W = -\boldsymbol{H}_j \cdot \boldsymbol{S}_j, \qquad (5)$$

is given by:

$$\boldsymbol{H}_j = Gb\boldsymbol{\epsilon} \cos \boldsymbol{q} \cdot \boldsymbol{R}_j. \qquad (6)$$

The wave vector $\boldsymbol{q}$ of the spin-density wave will be unrelated to interatomic distances of the lattice. On a statistical basis, half of the solute spins will be oriented parallel to $\boldsymbol{\epsilon}$ and the other half antiparallel.

The spin-density wave will constitute a collective, deformed state of the electron gas that is an eigenfunction of the Hartree–Fock equations with an energy greater than that of the ground state. SLATER[11] has pointed out that exchange potentials arising from a spin-density wave are attractive, and therefore tend to dynamically maintain the spin-density wave. This problem is formulated below, and complete dynamic stability of a spin-density wave is found to be possible for a particular magnitude of the wave vector $\boldsymbol{q}$, and then only if the mean exchange interaction between conduction electrons is sufficiently large. The spin-density wave can be thought of most easily by considering it to be the sum of two charge-density waves, one of electrons with spin parallel to $\boldsymbol{\epsilon}$ and the other of antiparallel spin, which are equal in magnitude but opposite in phase. Therefore, the total charge density of the gas remains uniform in space.

The higher energy of a spin-density wave precludes its formation in a pure metal. However, in the alloys under consideration, and at sufficiently low temperatures, the thermodynamic orientation of the solute spins in the effective field of the spin-density wave provides an interaction energy which more than compensates the higher energy of the spin-density wave. Therefore, below a critical temperature, the amplitude b of the wave will increase from zero and approach a maximum at 0°K.

Several appealing features of this mechanism can be appreciated immediately. The spin-density wave, together with its tendency to orient solute spins antiferromagnetically, vanishes above the critical temperature. Consequently, the residual, short-range solute–solute interactions of the direct overlap or pseudo-exchange type remain and can account for a positive paramagnetic Curie point, although they could never produce long-range ferromagnetic order. This mechanism provides, therefore, an easy explanation of how an antiferromagnet can be paramagnetically "ferromagnetic".

Since the wave vector of the spin-density wave

is not an integral multiple of a reciprocal lattice vector, the solute spins will see differing effective fields, of magnitude varying from 0 to Gb, depending on the location of each. Such a distribution of effective fields explains why the low-temperature specific heat arising from the coöperative interaction is linearly proportional to T far enough below the transition temperature.[5] The reason is that each incremental increase in temperature allows a new group of solute spins to reach the temperature at which they become thermally disoriented, and the integrated heat capacity of such a group is proportional to the temperature at which disorientation occurs.

VOL. 13 1960

One of the most unusual properties of copper–manganese alloys is the behavior of the remanent magnetization.[4] If a sample is cooled through the transition region in the absence of an applied field, little or no remanence occurs. A saturation remanence corresponding to about 1 atomic per cent alignment of the manganese spins can be induced by temporarily applying a field greater than 15 kOe. The remarkable feature is that the saturation remanence, once produced, can be reversed by fields of only a few kOe. Such behavior can be understood within the framework of the present model. It is necessary, of course, to have a domain structure; otherwise remanence could not occur on a macroscopic basis, since for a given domain the numbers of parallel and antiparallel solute spins can be expected to differ only by the square root of their number, according to statistical fluctuations. The smaller the domain size, the greater the relative unbalanced moment. The inhomogeneous distribution of solute atoms, to which the lack of sharpness of the transition region has already been attributed, should provide also a spontaneous means of domain formation.

Consider then an individual domain. Only the magnitude of the wave vector of the spin-density wave is fixed by the condition of self-consistency. Its direction and phase are free to adjust so as to minimize the free energy. In the presence of an applied field such adjustment will maximize the unbalanced moment parallel to the field. Presumably 15 kOe is sufficient to overcome all barriers which inhibit such adjustment, and which serve also to lock in the unbalance. Furthermore the axis of spin quantization assumed for the spin-density wave plays no role in the energetic or dynamic stability of the wave. Consequently all spin directions are free to rotate coherently, together with any unbalanced moment. Presumably the few-kOe field necessary for reversal provides opposition to magnetic dipole–dipole interactions or crystalline anisotropy arising from spin-orbit interactions, which have not been introduced into the model for the sake of simplicity. The foregoing discussion is admittedly conjectural, but it illustrates the ease with which the experimental phenomena can be interpreted.

One further enigma deserves discussion. Both copper–cobalt and copper–manganese alloys are paramagnetic, but only the latter undergo a coöperative transition at low temperatures.[4] Why is there a difference? A somewhat speculative explanation can be given in terms of the present model. As has already been mentioned, dynamic stability of a spin-density wave is possible only if the mean exchange interaction between conduction electrons of the metal is sufficiently large. On this basis alone, copper–cobalt and copper–manganese alloys should be similar. But the coöperative mechanism requires a further condition, specific to the solute: the paramagnetic orbital of the solute must be localized in a region small compared to the wavelength of the spin-density wave. (This wavelength is expected to be of the order of a lattice constant.) The physical origin of such a condition is clear. If the paramagnetic orbital, although localized, were to extend over several half-wavelengths of the spin-density wave, the orbital's exchange interaction with the wave would average to zero, and the effective fields necessary for long-range order would vanish. Since cobalt is two atomic numbers closer to copper than manganese it is to be expected that d-hole orbitals of cobalt in copper are less localized than those of manganese in copper. (Were the solute atomic number two units higher than cobalt, a d-hole orbital would be completely non-localized.) Although copper–cobalt alloys are paramagnetic, they show marked deviations from a Curie–Weiss law with regard to both field-dependence and concentration-dependence of the susceptibility.[4] This behavior, occurring even in alloys with ~ 1 atomic per cent cobalt, is indicative of long-range solute–solute interactions. In view of the screening action proved by HART[8], these interactions probably arise by direct overlap of extensive orbitals. If

such is indeed the case, the present theory explains also the basic difference between copper–cobalt and copper–manganese alloys.

2. DYNAMICS OF A SPIN-DENSITY WAVE

The free-electron model of a metal will be used in treating the quantum mechanics of a spin-density wave. (Such an approximation should not entail any essential compromise.) The objective is to find solutions of the Hartree–Fock equations for which the expectation value, (4), of the total spin-density operator is given by (3). The Hartree–Fock Hamiltonian for the problem is:

$$H = H_0 + A, \tag{7}$$

where H_0 is the electronic kinetic energy and A is the exchange operator, defined as follows:

$$A\psi(\mathbf{r}_1) \equiv -\sum_k [\int \phi_k^\dagger(\mathbf{r}_2)\psi(\mathbf{r}_2)V(r_{12})\, d^3r_2]\phi_k(\mathbf{r}_1). \tag{8}$$

Here, the set of (two-component) functions $\phi_k(\mathbf{r})$ are the occupied one-particle states of the electron gas, $V(r_{12})$ is the electron–electron interaction and ψ is any one-electron wave function in Hilbert space. It must be appreciated that A is a linear Hermitian operator under these general assumptions. It is not necessary to include the Hartree term in (7), since the deformations of the electron gas under consideration do not alter the total charge density, which is zero.

The usual solutions of (7) are plane waves. The solutions of interest here will be slightly perturbed plane waves. The amplitude b of the spin-density wave will be very small compared to unity. Consequently, it should be adequate to represent the perturbed plane waves as power series in b (retaining, of course, only the first-order correction terms). The following matrix elements of the exchange operator between plane waves may be defined:

$$\left.\begin{aligned} g^{\pm}(k) &\equiv \{\exp[i(k\pm q)\cdot r]\alpha|A|\exp(ik\cdot r)\alpha\} \\ h^{\pm}(k) &\equiv \{\exp[i(k\pm q)\cdot r]\beta|A|\exp(ik\cdot r)\beta\} \end{aligned}\right\} \tag{9}$$

where α and β are spin-up and spin-down spin functions relative to an arbitrary direction $\boldsymbol{\epsilon}$ of spin quantization. It is easy to see from (8) that there are no matrix elements of A between waves of opposite spin states if the occupied states, ϕ_k, have unmixed spin states, as will be assumed. (Such an assumption is therefore a consistent one.) The spin-up perturbed states can be written, for example:

$$\phi_k(r) = \exp(i\mathbf{k}\cdot\mathbf{r})\alpha\left(1+\frac{g^+(\mathbf{k})}{E_k - E_{k+q}}\exp(i\mathbf{q}\cdot\mathbf{r}) + \right.$$
$$\left. + \frac{g^-(\mathbf{k})}{E_k - E_{k-q}}\exp(-i\mathbf{q}\cdot\mathbf{r})\right), \tag{10}$$

where E_k is the single-particle kinetic energy. If these wave functions are used for the occupied states appearing in the exchange operator, one can evaluate the matrix elements (9) and thereby obtain a pair of coupled integral equations for $g^{\pm}(k)$. They are:

$$\left.\begin{aligned} g^+(K) &= -\sum_k\left[\frac{J(K-k)}{E_k - E_{k+q}}g^+(k) + \right. \\ &\qquad \left. + \frac{J(K-k+q)}{E_k - E_{k-q}}g^-(k)^*\right] \\ g^-(K) &= -\sum_k\left[\frac{J(K-k-q)}{E_k - E_{k+q}}g^+(k)^* + \right. \\ &\qquad \left. + \frac{J(K-k)}{E_k - E_{k-q}}g^-(k)\right] \end{aligned}\right\} \tag{11}$$

where $J(q)$ is the exchange integral associated with the electron–electron interaction:

$$J(q) \equiv \Omega^{-1}\int V(r)\exp(i\mathbf{q}\cdot\mathbf{r})\, d^3r, \tag{12}$$

Ω being the volume in which the wave functions are normalized. An identical set of equations applies to $h^{\pm}(k)$, but the requirement that the charge-density deviations of spin-up electrons cancel those of spin-down entails that the solutions of interest obey

$$h^{\pm}(k) = -g^{\pm}(k), \tag{13}$$

a relation that can always be satisfied by the solutions of equations (11) as a consequence of their partial linearity. (A solution multiplied by a *real* constant is also a solution.)

The factors in the integral equations have the following important symmetry properties:

$$\left.\begin{aligned} J(q) &= J(-q) \\ E_k &= E_{-k} \end{aligned}\right\} \tag{14}$$

As a consequence, it is relatively easy to see that if

$$f = [g^+(K), g^-(K)] \tag{15}$$

is a solution of (11), then:

$$f' = Tf \equiv [g^-(-K)^*, g^+(-K)^*] \tag{16}$$

is also a solution which is either $\pm f$ or is linearly independent of f. In the latter case it follows that the functions

$$F = [g^+(K)+g^-(-K)^*, g^-(K)+g^+(-K)^*] \tag{17}$$

and

$$F' = [g^+(K)-g^-(-K)^*, g^-(K)-g^+(-K)^*] \tag{18}$$

are also solutions, and which, when considered as basis vectors, span the same space as f and f'. But, $TF = F$ and $TF' = -F'$. Therefore it has been proved that basis vectors which span the space of solutions of equation (11) can *all* be chosen to satisfy:

$$f = \pm Tf \tag{19}$$

Basis vectors obeying the $+$ sign in (19) can then be written:

$$f = [g(K), g(-K)^*], \tag{20}$$

where $g(K)$ is the solution of the integral equation:

$$g(K) = -\sum_k [J(K-k)+J(K+k+q)]\frac{g(k)}{E_k - E_{k+q}}; \tag{21}$$

and basis vectors obeying the $-$ sign in (19) can be written:

$$f = [g(K), -g(-K)^*], \tag{22}$$

where $g(K)$ satisfies:

$$g(K) = -\sum_k [J(K-k)-J(K+k+q)]\frac{g(k)}{E_k - E_{k+q}}. \tag{23}$$

(Since the kernels of equations (21) and (23) are real, the basic functions of their solutions can be chosen real without loss of generality.) Each solution of (21) or (23) generates two linearly independent solutions of (11). For example, if $g(K)$ is a solution of (21), the two corresponding solutions of (11) are (20) and

$$f = [ig(K), -ig(-K)^*]. \tag{24}$$

Spin-density waves corresponding to (20) and (24) have the same structure, but are shifted in real space by one-quarter wavelength. The phase of a spin-density wave can be varied continuously, therefore, by appropriately varying the linear combination of these two degenerate solutions.

It has been proved that the space of solutions of the coupled integral equations, (11), is identical to the space generated by the solutions of the uncoupled equations (21) and (23). It has not been proved that solutions other than $g(K) \equiv 0$ exist. It should be noted that the wave vector $\boldsymbol{q}$ is a parameter in the integral equation, and its magnitude may play the role of an eigenvalue parameter. Treating equation (21) or (23) exactly is a very difficult task because of the complexity of the kernels. It seems necessary to make simplifying approximations in order to gain insight about the solutions. A very effective approximation is to neglect the variation of $J(q)$ by replacing it with a constant. This is perhaps a better approximation than it may seem at first sight, since much of the variation of $J(q)$ is averaged out by the integrations over the planes of constant-energy denominator. For a coulomb interaction

$$J(q) = 4\pi e^2/\Omega q^2. \tag{25}$$

The wave-number difference q appearing in the denominator of (25) will generally have a magnitude of the order of the radius k_0 of the Fermi sphere. Consequently, an average exchange integral, $\bar{J}$, can be defined:

$$\bar{J} = 4\pi e^2/\Omega k_0^2 \xi, \tag{26}$$

where ξ is the required numerical factor, of unit order of magnitude, in order that (26) be the appropriate average value of J as defined by equation (21). The total exchange energy per electron of a Fermi gas is:

$$-J = -9\pi n e^2/4k_0^2, \tag{27}$$

where n is the number of electrons per cm^3. It will be convenient to express $\bar{J}$ in terms of J:

$$\bar{J} = 16J/9n\Omega\xi. \tag{28}$$

If the constant $\bar{J}$ is inserted into (23), only the trivial solution, $g(K) \equiv 0$, results. If $J(q)$ is replaced by $\bar{J}$ in (21), a solution, $g(K) = g$, an arbitrary constant, is obtained provided that the

following equation can be satisfied for some value of q:

$$1 = -2\bar{J}\sum_k (E_k - E_{k+q})^{-1}. \tag{29}$$

If no root of this equation exists, only the trivial solution ($g = 0$) is possible, and a spin-density wave cannot occur. The integrand in equation (29) has singularities, but one can prove that the principal value of the integral is the correct interpretation. The result is,

$$\sum_k (E_k - E_{k+q})^{-1} = -(3n\Omega/8E_F)p(q/2k_0), \tag{30}$$

where E_F is the Fermi energy and $p(x)$ is given by:

$$p(x) = \frac{1}{2} + \frac{1-x^2}{4x}\log\left|\frac{1+x}{1-x}\right|. \tag{31}$$

If the foregoing equations are combined, the relation determining q is:

$$1 = (4J/3\xi E_F)p(q/2k_0). \tag{32}$$

Since $p(0) = 1$ and is monotonic, decreasing for increasing x, (32) has a solution if

$$J > (3\xi/4)E_F. \tag{33}$$

This inequality is not too stringent, since $J \sim E_F$ for most metals. Dynamic stability of a static spin-density wave, at least in some metals, seems to be an assured possibility.

The amplitude b (see equation (3)) of the spin-density wave can be computed directly from the operator in (4) and the wave functions (10). One finds

$$b = -(9\xi\eta/8J)g, \tag{34}$$

where $\eta = n/N$, the number of conduction electrons per atom, and where use has been made of equations (30) and (32). The $-$ sign in (34) reflects the attractive character of exchange interactions. Wherever the spin density is positive, the exchange potential is negative (as is necessary for a dynamically stable situation).

The remaining task of this section is to calculate the energy increase associated with a spin-density wave. This can be accomplished by summing the kinetic-energy change of each occupied state together with one-half the energy change arising from the exchange potential. Employing the wave functions (10) and taking care to normalize, then to order g^2, one finds the kinetic energy increase per cm^3 to be,

$$\Delta T = (9\xi n/8J)g^2. \tag{35}$$

Similarly, the total coulomb energy change turns out to be the negative of (35). Consequently, in the Hartree–Fock approximation, the energy of a spin-density wave is zero. It is necessary to include correlation-energy changes in order to obtain an energy increase. The magnitude of the correlation-energy change can be estimated by means of the following relatively crude argument. Consider a small region in the Fermi gas for which the number of $+$ spins is $M(1+\epsilon)$. The exchange energy is proportional to the number of pairs of parallel spins, which is:

$$(1/2)M^2(1+\epsilon)^2 + (1/2)M^2(1-\epsilon)^2 = M^2(1+\epsilon^2). \tag{36}$$

The correlation energy is proportional (mainly) to the number of pairs of antiparallel spins, which is:

$$M(1+\epsilon)M(1-\epsilon) = M^2(1-\epsilon^2). \tag{37}$$

From equations (36) and (37) it is clear that the fractional (algebraic) decrease in the exchange energy should roughly equal the fractional increase in the correlation energy, arising from a small spin-density deviation. Therefore, if we let $-C$ be the correlation energy per electron of the Fermi gas, the energy W per cm^3 of a spin-density wave is approximately C/J times the kinetic-energy increase (35). With the help of (34), this energy density can be written:

$$W(b) = 8nCb^2/9\xi\eta^2. \tag{38}$$

For typical metals the correlation energy per electron, C, is ~ 1 eV.[12]

Equation (38) is the only result of the present section that is needed for treating the thermodynamics of the antiferromagnetic phase. However, the main burden of the foregoing development has been to give theoretical justification to static spin-density waves as a mechanism of long-range antiferromagnetic order.

3. THERMODYNAMICS OF THE ANTIFERROMAGNETIC PHASE

The magnetic contributions to the free energy of a dilute alloy can be evaluated immediately by

considering the distribution of effective fields, (6), experienced by the paramagnetic solute spins as a result of their interaction, (5), with the spin-density wave. If the orbitals of the solute spins are sufficiently localized, compared to a wavelength of the spin-density wave, the probability distribution for the (absolute) magnitude of the effective field H is:

$$P(|H|) = (2/\pi)(G^2b^2-H^2)^{-1/2}, \tag{39}$$

for $|H|$ smaller than the maximum value, Gb. The partition function of a spin S_j in a field H_j is:

$$Z_j = \sinh\{\beta[S_j+(1/2)]H_j\}/\sinh(1/2)\beta H_j, \tag{40}$$

where $\beta = 1/kT$. The free energy of such a spin is:

$$A_j = -(\log Z_j)/\beta. \tag{41}$$

The magnetic free-energy density of the dilute alloy can be expressed as the average of (41), weighted by the probability distribution (39), and added to the self-energy of the spin-density wave:

$$A = W(b) - \frac{2cN}{\pi\beta}\int_0^{Gb}(G^2b^2-H^2)^{-1/2}\times \times \log\left[\frac{\sinh\beta[S+(1/2)]H}{\sinh(1/2)\beta H}\right]dH, \tag{42}$$

where c is the atomic fraction of solute spins. The amplitude b of the spin-density wave must be chosen so as to minimize (42). $b(T)$ will have a maximum value $b(0)$ at $T = 0$, and will become zero at the transition temperature T_c. By appropriate power-series expansions of (42) near $T = 0$ and T_c, it is easy to derive the following results: near $T = 0$, including terms up to order T^2, (42) becomes:

$$A = \frac{8nCb^2}{9\xi\eta^2} - \frac{2cNSGb}{\pi} - \frac{2\pi ScN}{3(2S-1)Gb}k^2T^2. \tag{43}$$

The value of b which minimizes this free energy is, to order T^2,

$$b(T) = \frac{9\xi\eta SG}{8\pi C}c - \frac{8\pi^3C}{27\xi\eta S(2S+1)G^3c}k^2T^2. \tag{44}$$

The low-temperature free energy is determined by inserting (44) into (43):

$$A = -\frac{9n\xi S^2G^2c^2}{8\pi^2C} - \frac{16\pi^2NC}{27\xi\eta(2S+1)G^2}k^2T^2. \tag{45}$$

Since the specific heat (at constant volume) is given by:

$$C_V = -T\left(\frac{\partial^2 A}{\partial T^2}\right)_V, \tag{46}$$

the low-temperature magnetic specific heat is

$$C_{VM} = \gamma_M T, \tag{47}$$

where,

$$\gamma_M = \frac{32\pi^2NCk^2}{27\xi\eta(2S+1)G^2}. \tag{48}$$

This is a most remarkable result, because γ_M is *independent* of solute concentration. The experimental measurements of ZIMMERMAN[5] are in complete accord with this conclusion. He finds

$$\gamma_M \sim 4\times10^{-3}\ J/\text{mole, deg}^2, \tag{49}$$

for samples of 1/6, 1/2, 1, 2, and 4 atomic per cent manganese in copper. The magnitude of (49), as well as the lack of concentration-dependence, is in excellent agreement with (48). (One should appreciate that the maximum temperature below which (47) is valid is proportional to the solute concentration.)

By appropriate power-series expansions near T_c, the following results can be established:

$$kT_c = \frac{3\xi\eta S(S+1)G^2}{32C}c. \tag{50}$$

The critical temperature is proportional to the solute concentration. By combining (48) and (50), all of the less-known parameters can be eliminated:

$$\gamma_M T_c = \frac{\pi^2}{9}\cdot\frac{S(S+1)}{2S+1}cNk. \tag{51}$$

If the experimental value (49) for γ_M (together with $S = 2$) is used in this expression, the indicated critical temperature of a 1 atomic per cent manganese alloy is $T_c = 27°\text{K}$. The temperature variation of b, the spin-density wave amplitude, near T_c is given by:

$$b(T) \sim (T_c-T)^{1/2}, \tag{52}$$

VOL. 13 1960

and $b = 0$ for $T > T_c$, as would be expected. The variation of b with T throughout the antiferromagnetic phase can, with an accuracy of 2 per cent or better, be taken to be:

$$b(T) = b(0)[1-(T/T_c)^2]^{1/2}, \quad (53)$$

where $b(0)$ is given by the first term in (44).

The magnetic contribution to the specific heat just below T_c is:

$$C_{VM}(T_c) = \left[\frac{30(2S+1)}{\pi^2(2S^2+2S+1)}\right]\gamma_M T_c. \quad (54)$$

Since the coefficient in brackets has a value 1·17 (for $S = 2$), the variation of the specific heat throughout the antiferromagnetic region deviates very little from the linear behavior given by equation (47). The magnetic specific heat rises slightly above the linear low as T_c is approached, and then drops discontinuously to 0 above T_c. This sudden drop is not very sharp experimentally, especially in the less-dilute alloys. Such broadening is probably associated with inhomogeneities, and one may expect a high-temperature tail in the magnetic specific heat as a result of solute–solute pseudo-exchange interactions, particularly in the more concentrated alloys. The temperature at which C_{VM}/T drops to half its low-temperature value should provide a relatively reliable measure of T_c. The experimental results for 1/6 and 1/2 atomic per cent manganese alloys, so interpreted, yield:

$$T_c = 26\pm4^\circ\text{K/atomic per cent}, \quad (55)$$

a value in good agreement with equation (51), and with the onset of the anomalous resistivity decrease.[3] If $\xi, \eta \sim 1$, and $C \sim 1$ eV, then the s–d exchange interaction parameter G which yields agreement between (50) and (55) is:

$$G = 0{\cdot}6 \text{ eV}. \quad (56)$$

This value also gives agreement between the theoretical and experimental magnitudes of the anomolous resistivity decrease,[10] and compares well with the estimate, (2), derived from spectroscopic data of free ions.

4. CONCLUDING REMARKS

The applications of the foregoing theory to the electrical and magnetic properties of copper–manganese alloys will be elaborated in a subsequent paper. The static spin-density wave mechanism of antiferromagnetic order may obtain in systems other than dilute paramagnetic alloys. For example, it may provide a mechanism of nuclear antiferromagnetism, the hyperfine interaction playing the role of the s–d exchange interaction for coupling localized spins to the spin-density wave. The mechanism may also operate in a pure metal, such as α-manganese, which is antiferromagnetic and which has an anomalously large linear contribution to the low-temperature specific heat.

Acknowledgement—The author is grateful to Dr. A. ARROTT for many helpful discussions regarding magnetic structures and behavior.

REFERENCES

1. OWEN J., BROWNE M., KNIGHT W. D. and KITTEL C., *Phys. Rev.* **102,** 1501 (1956).
2. OWEN J., BROWNE M., ARP V. and KIP A. F., *J. Phys. Chem. Solids* **2,** 85 (1957).
3. SCHMITT R. W. and JACOBS I. S., *J. Phys. Chem. Solids* **3,** 324 (1957).
4. JACOBS I. S. and SCHMITT R. W., *Phys. Rev.* **113,** 459 (1959).
5. ZIMMERMAN J. E. To be published.
6. GORTER C. J., VAN DEN BERG G. J. and DE NOBEL J., *Canad. J. Phys.* **34,** 1281 (1956).
7. SATO H., ARROTT A. and KIKUCHI R., *J. Phys. Chem. Solids* **10,** 19 (1959).
8. HART E. W., *Phys. Rev.* **106,** 467 (1957).
9. YOSIDA K., *Phys. Rev.* **106,** 893 (1957).
10. YOSIDA K., *Phys. Rev.* **107,** 396 (1957).
11. SLATER J. C., *Phys. Rev.* **82,** 539 (1951).
12. SEITZ F., *Modern Theory of Solids* Chap X. McGraw-Hill, New York (1948).

APPENDIX

Consideration of objections

In presenting the idea of a static spin density wave, the author has encountered frequently several questions related to the theoretical framework within which the concepts associated with such excited states have been elaborated. It seems worthwhile to discuss in greater detail some of these "objections".

Objection I

The criterion, equation (33), for the occurrence of static spin density waves as collective excited states of a Fermi gas is similar in form to that for free electron ferromagnetism, ($J > 1{\cdot}36E_F$, in the Hartree–Fock approximation). If one replaces the Coulomb interaction by a delta function, the resulting criteria are in fact

numerically identical. Consequently, it would appear that spin density waves can occur only if the Fermi gas were ferromagnetically unstable, which fact would preclude their formation in the first place (at least for the proposed applications).

Reply to objection 1

The factor ξ in equation (33), and hence the precise criterion for a spin density wave, can be determined only by solving the integral equation (21) exactly. The value of ξ is related to the average value of the exchange integral (equation 25) according to the averaging process defined by the integral operator in equation (21). The best *a priori* estimate of ξ, perhaps, is the value required for the approximation (26) to yield the correct total exchange energy, which is known exactly and which defines thereby *an* average exchange integral. Such an estimate of ξ yields 4/9; so that the condition for static spin density waves would be, $J > E_F/3$. This criterion is weaker than that for ferromagnetism by at least a factor of 4, and would allow a Fermi gas to have static spin density waves as long as the electron concentration were between 1 and 64 times the limiting density compatible with ferromagnetism. This result provides, therefore, an adequate "margin of safety" and justifies the concept of static spin density waves as a working hypothesis, subject of course to subsequent experimental or theoretical verification. Lack of precision in the latter respect cannot be considered serious, however, since by analogy, one should remember that the entire electron theory of metals rests on the theoretically *unproved* assumption (among others) that the Hartree–Fock ground state of an ideal Fermi gas is the familiar Fermi sphere of occupied momentum states.

Objection 2

The proposed mechanism of antiferromagnetism is similar to the mechanism of ferromagnetism suggested by ZENER. However, YOSIDA has shown (with simplifying approximations) that the energy contribution resulting from first order excitation of ferromagnetic excited states of the Fermi gas is cancelled by second order contributions of the *s–d* exchange interaction. Similarly, one might expect such second order contributions to cancel as well the antiferromagnetic interaction resulting from first order excitation of a static spin density wave.

Reply to objection 2

No such cancellation occurs in this case. The totality of intermediate states employed by YOSIDA in his second order calculation exhausts all of Hilbert space. No further second order contributions remain to be evaluated. (Of course, in so far as intermediate states employed in a perturbation calculation should approximate eigenstates, the recognition of collective modes—e.g. plasma waves, spin density waves, etc.—as better "eigenstates" than the single particle excitations, of which they are linear combinations, will suggest improvements to such computations. But the improvements would involve replacement rather than addition of terms.) The spin density wave mechanism of antiferromagnetism assumes and incorporates YOSIDA's work as a foundation. The second order *s–d* exchange interaction between solute spins is presumed to be always present (and to be responsible for the positive paramagnetic Curie temperature of Cu–Mn alloys). A localized spin which interacts with a spin density wave is not just the "bare" solute ion spin-orbital, but that together with its localized conduction electron, spin density readjustment, as envisioned by YOSIDA. It is incorrect to think that spin density waves give rise to a solute–solute interaction of infinite range. Rather, the spin density wave must be considered as having independent existence, and as interacting individually with localized solute spins; the very *tendency* towards long range antiferromagnetic order vanishes above the Néel point, where the spin density wave amplitude is zero.

Objection 3

It is difficult to understand why a single spin density wave having a unique wave vector magnitude, q_0, should be involved so pre-eminently instead of a cooperative participation by many spin density waves with a distribution in q; and it seems odd that the allowed amplitudes of the spin density wave form a continuous distribution rather than a discrete set like the harmonic oscillator levels of an ordinary collective mode.

Reply to objection 3

The special character of a static spin density wave is that it is *static*. The uniqueness of the eigenvalue q_0 is associated with the employment of stationary state perturbation theory in deriving the integral equations (11). There are, to be sure, non-static spin density waves, which oscillate harmonically in time, and these can be constructed by employing time dependent perturbation theory in a straightforward fashion. One finds that the frequency ω of such waves varies with q according to $(q_0-q)^{1/2}$. (These collective modes will be quantized and will yield a negligible contribution, $\sim Nk(kT/E_F)^2$, to the specific heat.) The molecular fields arising from such waves will oscillate in time, and that is why they are unimportant for long range magnetic order. It seems likely, however, that waves with q near q_0 may play a role in the description of antiferromagnetic domains. The Hamiltonian of a collective mode can be written, $P^2+\omega^2Q^2$, where P and Q are appropriate generalized coördinates. It is well known that a Hamiltonian of this type leads to discrete energy levels, but only as long as $\omega \neq 0$. When $\omega = 0$, only the P^2 term remains, and the allowed energies form a continuous spectrum, analogous to that of free particles.

Objection 4

Since solutions of the integral equation (11) obey a superposition theorem, it seems that a spiral spin density wave (such as a cos qz polarized in x plus a sin qz polarized in y) could be constructed and would lead to an antiferromagnetic state of lower energy, since the spin

VOL. 13 1960

density would have the same magnitude everywhere, changing only its direction with position.

Reply to objection 4

The integral equations satisfy a superposition principle, but only *after* a specific axis of spin quantization has been chosen. One can superpose two spin density waves only if they have the same axis of spin quantization. To construct a spiral spin density wave as suggested, one would require conduction electron spin states simultaneously quantized along two perpendicular axes—an obvious impossibility.

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 13. pp. 81–87. Printed in Great Britain.

ON THE INFRARED EMISSION IN ZnS:Cu—EFFECT OF SULFUR PRESSURE AND ALUMINUM*

E. F. APPLE and J. S. PRENER

General Electric Research Laboratory, Schenectady, New York

(*Received* 3 *July* 1959; *revised* 16 *November* 1959)

VOL. 13 1960

Abstract—The intensity of the infrared emission in ZnS:Cu phosphors has been studied as a function of the sulfur pressure under which the phosphors were heated at high temperatures and of Al concentration in the phosphors. The observed dependence is interpreted using a model in which the infrared absorption and emission arises from hole transitions between states of a normally unionized Cu acceptor and the valence band, and in which sulfur vacancies behave as doubly ionizable donors in ZnS. The effect of Al impurities also follows from the model.

1. INTRODUCTION

INFRARED emission bands peaking at 1·5μ and 1·65μ in ZnS phosphors were observed by DUMBLETON and GARLICK[1] in 1954. In an investigation of this infrared emission in 1956, BROWNE[2] found that ZnS phosphors could be divided into two classes depending on the conditions necessary to excite this infrared emission.

In his Class I phosphors, the infrared emission could be observed by exciting in bands peaking at 0·76μ and 1·35μ and without previous or simultaneous ultraviolet excitation. His Class II phosphors required simultaneous excitation by the infrared and 0·365μ radiation so that under these conditions of excitation the phosphors showed visible as well as infrared emission. Browne stated that the Class I phosphors were nonstoichiometric in which the only appreciable impurity could have been chlorine and that only the Class II phosphors contained copper as an impurity. However, MEIJER[3] and HALSTED *et al*[4] showed that the aforementioned infrared absorption and emission were observed only in those ZnS samples containing copper as an impurity.

The purpose of the investigation reported in this paper was to define the conditions of preparation of copper activated ZnS phosphors which showed the infrared emission under infrared excitation alone. Specifically the subject of this paper covers the effect on the intensity of the infrared emission of variations in sulfur pressure in equilibrium at high temperatures with the phosphor, ZnS:Cu, and of the concentration of Al donor impurities.

2. EXPERIMENTAL

(a) *Preparation of samples*

Cu and Al were added as solutions of the sulfates to a water slurry of ZnS (General Electric Co., luminescent grade) which previously had been heated in a stream of purified H_2S† successively at 400°C for 2 hr, 800°C for 2 hr and 1150°C for 3 hr, and which was nonluminescent under 3650 Å excitation even at −196°C. After drying at 100°C, the phosphor mix was heated at 1150°C in an H_2S stream for 3 hr.

A sample of the fired phosphor weighing 3–4 g was placed in the top part of the quartz firing tube shown in Fig. 1 and 0·2–1·8 g S‡ was added to the bottom part of the tube. The tube was evacuated

* This paper was presented at the 1959 Spring Meeting of the Electrochemical Society held in Philadelphia, Pa.

† The H_2S (Matheson Co.) used was bubbled through saturated $Ba(OH)_2$ solution, then passed through silica gel, P_2O_5, and a cold trap held at −50°C before coming in contact with the sample.

‡ The sulfur used was obtained by decomposition of purified H_2S. Prior to use, it was fused under vacuum to get rid of volatile inclusions.

to about 3×10^{-5} mm Hg, sealed off and placed in the three-zone furnace (shown in Fig. 1) which was designed so that the ZnS phosphor sample and the sulfur reservoir could be held at constant

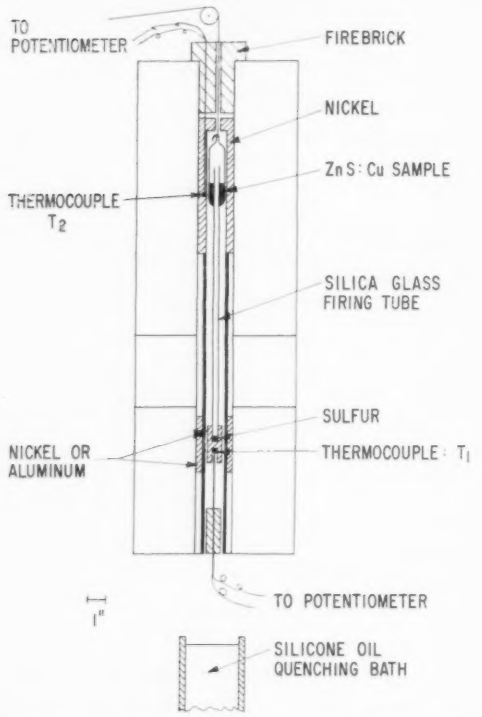


FIG. 1. Furnace and silica glass tube for firing under controlled sulfur pressure.

temperatures in the ranges 700–1100°C and 125–830°C, respectively. The temperature control was better than ± 2°C. In this way the vapor pressure of the sulfur over the sample was determined by the temperature of the sulfur reservoir.* The vapor pressure of sulfur in all the experiments exceeded the pressure of sulfur due to the dissociation of ZnS at 940°C.[5] After firing for 3 hr which time was sufficient to reach equilibrium (see Fig. 2), the tube was quenched in silicone oil. The quenching procedure was necessary to insure reproducibility between samples fired under the same conditions and also to give maximum i.r. effects, as is also shown in Fig. 2. X-ray diffraction studies showed that all samples had the hexagonal zinc sulfide structure.

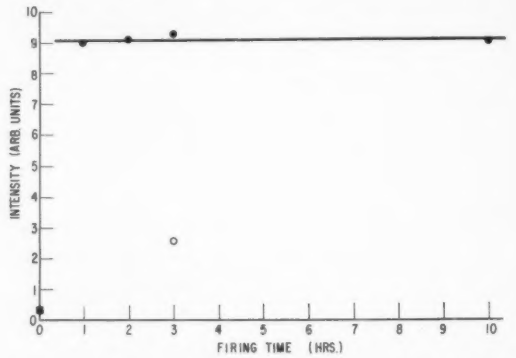


FIG. 2. Infrared emission intensity at −196°C of ZnS: 5×10^{-5}Cu as a function of firing time and quenching rate. Samples were equilibrated at 940°C under a sulfur pressure of $1{\cdot}1 \times 10^{-2}$ atm. ● quenched; ○ air cooled; ■ H_2S fired only.

(b) *I.R. detection system*

The resulting phosphors equilibrated under various sulfur pressures were excited at −196°C by a band extending from about 0·7 to 1·0μ isolated from a tungsten lamp with a Corning No. 2600 filter and a 1-cm H_2O cell. The infrared emission of the phosphor was passed through Corning No. 9788 and 2540 filters which absorbed reflected exciting radiation, and was then detected by a lead sulfide cell (Kodak Ektron Detector, Type II) cooled to −78°C. The photocurrent was measured with the aid of a sensitive d.c. galvanometer and any drifts in the photocell sensitivity or intensity of exciting radiation were monitored by use of a "standard" diffuse reflector (*BN* powder covered by a diffusing quartz disc). The geometry was such that samples could be reproducibly placed in the same position with respect to the exciting source and detection system. All the phosphors had infrared emission bands peaking at 1·67μ (0·74 eV) and 1·55μ (0·80 eV)[4] whose positions and relative intensities were independent of the sulfur pressure under which the samples were fired. Measurements of the infrared emission intensity were carried out at −196°C since the intensity decreases with increasing temperature as already reported by BROWNE[2].

* Vapor pressure data for sulfur obtained from: J. R. WEST, *Indstr. Engng. Chem.* **42**, 713 (1950); K. K. KELLEY, *Bull. U.S. Bur. Min.* no. 383 p. 99 (1935).

VOL. 13 1960

3. RESULTS

(a) *Effect of sulfur pressure*

A series of ZnS samples containing from 10^{-6} to 5×10^{-4} g atoms Cu/mole ZnS were fired under 0·8 atmosphere sulfur pressure. As shown in Fig. 3 the i.r. emission intensity increases with Cu

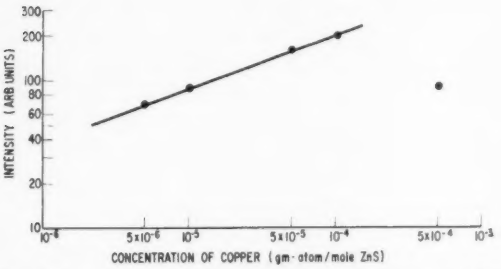


FIG. 3. Infrared emission intensity of ZnS:Cu at −196°C as a function of Cu concentration. Samples were equilibrated at 940°C under a sulfur pressure of 0·8 atm.

concentration in the range 10^{-6}–10^{-4}; concentration quenching causes the intensity to drop in the 5×10^{-4} sample. In one group of experiments, phosphors of a given composition were heated at 940°C under sulfur pressures ranging from 7×10^{-5} to 30 atm. The dependence of I.R. emission

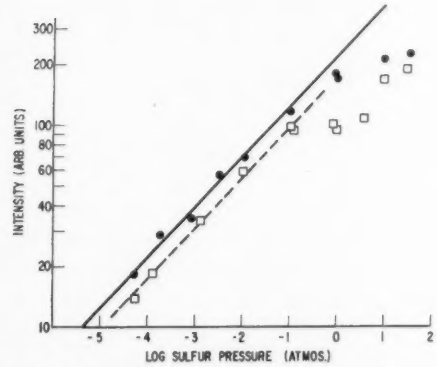


FIG. 4. Infrared emission intensity of ZnS:Cu at −196°C as a function of sulfur pressure. Samples were equilibrated at 940°C. Lines are drawn with a +1/4 slope. ● ZnS:5×10^{-5}Cu; □ ZnS:10^{-5}Cu.

intensity on sulfur pressure in ZnS:10^{-5} Cu and ZnS:5×10^{-5} Cu is shown in Fig. 4. In the range from 7×10^{-5} to about 10^{-1} atmospheres sulfur, the log of the infrared emission intensity increases linearly with the log of the sulfur pressure. The line shown is drawn with a +1/4 slope. Above 0·1 atm, the intensity levels off with increasing sulfur pressure. Similar results were obtained with ZnS:5×10^{-5} Cu, 10^{-5} Al as shown in Fig. 5.

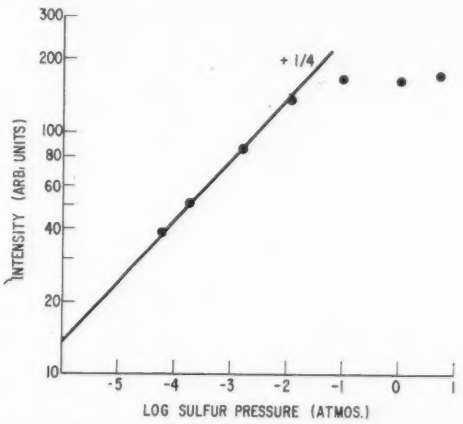


FIG. 5. Infrared emission intensity of ZnS:5×10^{-5} Cu, 10^{-5} Al at −196°C as a function of sulfur pressure. Samples were equilibrated at 940°C. Line is drawn with +1/4 slope.

(b) *Effect of donor concentration*

In another group of experiments, phosphors with 5×10^{-5} Cu and varying Al concentrations were heated at 940°C under 9 atm sulfur pressure

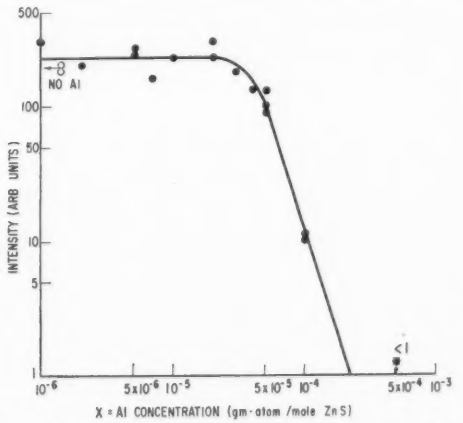


FIG. 6. Infrared emission intensity at −196°C of ZnS: 5×10^{-5}Cu, X Al as a function of Al concentration. Samples were equilibrated at 950°C under 9 atm of sulphur.

(in saturation region). The i.r intensity as a function of concentration is shown in Fig. 6. Over the range of 0 Al added to 3×10^{-5} Al added, the i.r emission intensity is essentially constant, from 3×10^{-5} to 5×10^{-5} Al the intensity drops by a factor of 2 and from 5×10^{-5} Al to 5×10^{-4} Al it drops by a factor of 100.

(c) *Reflectance measurements*

It was observed that the phosphors containing only Cu had a near-white body color when fired in H_2S but became grey when fired under the higher sulfur pressures. The greyness was not reduced by washing the phosphors in 10% KCN solution in which any copper sulfide present on the surface as a separate phase would be soluble. Reflectance measurements shown in Fig. 7 indicate the greyness is due to an absorption band extending across the visible part of the spectrum. In addition, there is a narrow absorption band peaking at $1{\cdot}3\mu$. The same reflectance spectrum was observed in a ZnS phosphor containing Cu produced by the radioactive decay of Zn^{65}.[6]

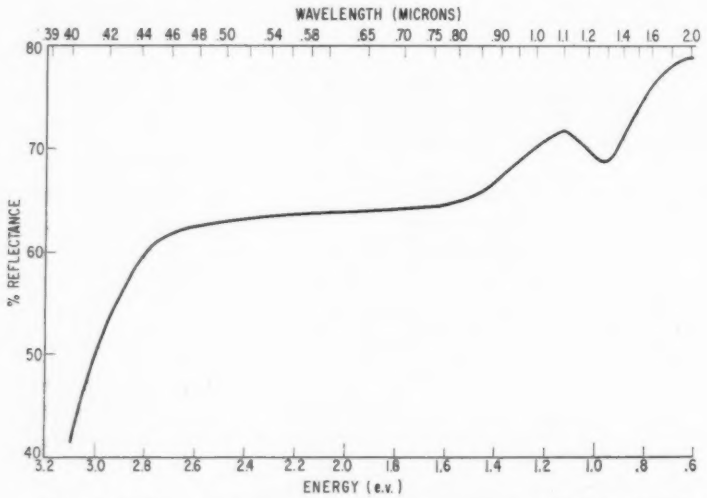


FIG. 7. Reflectance spectrum at room temperature of $ZnS:5\times10^{-5}$ Cu (reflectance of MgO taken as 100). Sample equilibrated at 940°C under 30 atm of sulfur.

(d) *Ultraviolet excited luminescence*

In the series of phosphors described in Section 3(b), the luminescence under 3650 Å excitation changed from orange-red to blue to green with increasing Al concentration as reported by AVEN and POTTER[7]. In the ZnS phosphors containing copper as the only added impurity described in Section 3(a), the luminescence under 3650 Å excitation consisted only of the orange-red band[7] in those samples equilibrated under low sulfur pressures. At the higher sulfur pressures a green band and a weaker blue band are observed in addition to the strong orange-red band. As the temperature is decreased from room temperature to that of liquid nitrogen, the ratio of the intensities of the green and blue bands to the orange-red band decreases. This ratio is also composition dependent in that the ratio increases with decreasing copper concentration.

4. DISCUSSION

The infrared emission and absorption observed in copper-activated ZnS phosphors has been interpreted as being due to hole transitions between the valence band and two localized states associated with the copper impurity[3, 4, 6] as shown in Fig. 8. From its position in the Periodic Table, it has been presumed that copper may act as a singly ionizable acceptor in ZnS giving rise to a ground and excited states for a hole with energy levels above the

valence band. With a hole in the ground state (upper level in Fig. 8), the copper is neutral with respect to the zinc for which it substitutes and can be represented by Cu°. With the hole ionized into the valence band so that the ground state and the excited hole state are occupied by electrons,

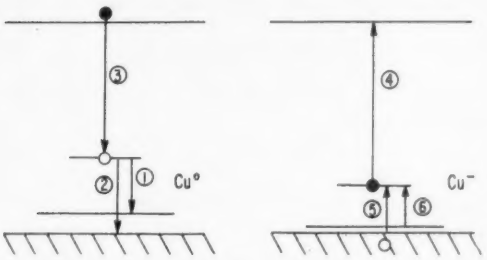


FIG. 8. Levels introduced by a copper substituting for Zn in ZnS. Hole transition (1): absorption band peaking at 0·95 eV. Hole transition (2): broad absorption starting at 1·3 eV. Electron transition (3): green emission band peaking at 2·4 eV. Electron transition (4): broad absorption starting at 2·9 eV. Hole transition (5): infrared emission band peaking at 0·80 eV. Hole transition (6): infrared emission band peaking at 0·74 eV. Band gap of ZnS:3·70 eV.

the copper is negatively charged and can be represented by Cu^-. In those phosphors which show emission in the infrared with bands peaking at 0·74 eV and 0·80 eV without previous ultraviolet excitation (BROWNE's Class I), the upper level is occupied in the dark by a hole. The sharp absorption band peaking at 0·95 eV corresponds to the excitation of the hole to an excited state and the broad absorption starting in the vicinity of 1·3 eV and extending to higher energies (see also Ref. 4) corresponds to ionization of the hole into the states of the valence band. The reverse transitions give rise to the two emission bands. On this model the strong absorption starting around 2·9 eV and extending to higher energies corresponds to the transition of an electron from the upper state due to copper to the states of the conduction band (see also Ref. 8). Finally the green luminescence peaking at 2·39 eV observed in copper activated ZnS phosphors arises from the recombination of a conduction electron (or one in a state very close to the conduction band) with a hole trapped at the copper acceptor.[9] The origin of the blue emission band observed in ZnS phosphors containing copper as impurity is not well understood at the present time. The phosphors showing infrared emission may exhibit orange-red, blue or green luminescence (or combinations of the three) under 3650 Å excitation (see Section 3d). However, the green and blue emissions are not observed in those phosphors containing copper as the only added impurity and equilibrated at the lower sulfur pressures even though they contain the same Cu levels as are present in those phosphors showing the green and blue emission bands. When there are two or more recombination processes competing for excitation energy, the observed luminescence depends not only on the presence of appropriate levels but also on the details of the recombination kinetics. The dependence on temperature and composition of the ratio of the green and blue intensities to the red intensity indicates that such competing processes are involved in ZnS:Cu phosphors. This was also shown in the studies of APPLE and WILLIAMS on a series of ZnS phosphors containing Cu and Ga or In impurities which may or may not show green luminescence depending on the excitation temperature and phosphor composition.[10] The difference in the level structure depending on whether the copper acceptor is ionized or not is due to lattice polarization effects and accounts for the observed shift between excitation and emission. It is significant that the various energy differences shown in Fig. 8 add up to the band gap energy of 3·7 eV at −196°C. Further evidence for hole transitions of the type described above is given in the work of DROPKIN and ULLMAN[11], in which they showed that excitation in the infrared absorption bands led to photoconductivity in ZnS single crystals. They presented evidence to show that the photocurrents were carried by holes. POTTER and AVEN[12] have shown by thermoelectric measurements that ZnS crystals having the infrared absorption and emission properties described above are *p*-type semiconductors. We shall, therefore, assume the correctness of the model for the origin of the infrared absorption and emission.

With different layers of polycrystalline ZnS phosphors of the same particle size distribution in which absorption of incident radiation is taking place in impurity centers, the power absorbed in the phosphor layers will be very nearly proportional to the power incident upon the sample and the concentration of absorbing species, provided

that the reflectance of the samples for the incident light is greater than about 0·6.[13] Even our most absorbing samples equilibrated under the highest sulfur pressures had reflectances greater than 0·6 in the infrared excitation region. Under these conditions the application of the model taken for the absorption and emission processes indicates that the emitted power should be proportional to the excitation intensity and to the concentration of Cu° provided that the former is not so high so as to saturate the absorbing centers. This would show up as a nonlinear relation between the excitation and fluorescent intensities. Measurements have shown a linear relationship over a wide range of excitation intensities.

The copper activated ZnS phosphors were equilibrated under various pressures of sulfur at high temperatures and quenched rapidly to room temperature. We assume that the atomic defects are frozen in during quenching at their equilibrium concentrations determined by the high temperature T and sulfur pressure p_{S_2}.[14] We have already indicated that copper can substitute for zinc in zinc sulfide either as Cu° or Cu^{-} and that the infrared emission intensity is proportional to the concentration of the former. The observed one fourth power dependence of the infrared emission intensity on the pressure of sulfur in the gas phase can be understood if four Cu° species are transformed into four Cu^{-} species when two sulfur atoms from the solid are transferred to the gas phase to form a sulfur molecule. We can write for the solid–gas equilibrium.

$$Zn\ Cu^{\circ}_{4\delta}\ S_{1+4\delta} \rightleftarrows Zn\ Cu^{-}_{4\delta}\ V^{++}_{2\delta}\ S_{1+2\delta} + \delta S_2 \quad (1)$$

V^{++} represents a sulfur vacancy which is doubly charged with respect to S in the solid.

Applying the law of mass-action with the condition that $\delta \ll 1$, we get

$$\frac{p_{S_2}[V^{++}]^2[Cu^{-}]^4}{[Cu^{\circ}]^4} = K \quad (2)$$

In addition we have the conditions of charge and mass balance in the solid.

$$[Cu^{-}] = 2\,[V^{++}] \quad (3)$$

$$[Cu^{\circ}]+[Cu^{-}] = [Cu]_{total} = \text{constant} \quad (4)$$

Equation (4) assumes, of course, that there is no precipitation of Cu from the solid solution. At low sulfur pressures, the equilibrium represented by equation (1) will be displaced to the right, so that the condition of mass balance can be approximated by

$$[Cu^{-}] = [Cu]_{total} \quad (5)$$

and equation (2) then leads to the result

$$[Cu^{\circ}] = \text{constant} \cdot p_{S_2}{}^{1/4} \quad (6)$$

At high sulfur pressure the mass balance equation is approximated by

$$[Cu^{\circ}] = \text{constant} \quad (7)$$

Thus the one-fourth power relationship at low sulfur pressures and the saturation observed at higher sulfur pressures can be explained. Since a sulfur vacancy can give up two electrons to a copper acceptor, it behaves as a double donor in ZnS with its two levels above that of the acceptor level introduced by Cu.

Aluminum impurities act in a similar manner since their position in the Periodic Table suggests that they can act as singly ionizable donors. The condition of charge balance becomes:

$$[Cu^{-}] = [Al^{+}]+2\,[V^{++}] \quad (8)$$

From measurements of thermal glow curves HOOGENSTRAATEN has found that Al introduces an electron trapping level with a thermal depth of about 0·25 eV below the conduction band.[15] In view of the assumed model for the origin of the infrared absorption and emission the presence of holes in the copper levels in the dark at low temperatures means that all the aluminum donors are ionized. Therefore:

$$[Al^{+}] = [Al]_{total} \quad (9)$$

Furthermore the experiments with aluminum doping were carried out at high sulfur pressures so that we may take

$$2[V^{++}] < [Al^{+}],$$

hence:

$$[Cu^{\circ}] = [Cu]_{total} - [Al]_{total} \quad (10)$$

When the aluminum concentration is much smaller than the copper concentration, the infrared intensity will be almost constant. However, when the concentration of added aluminum

approaches that of the copper, the infrared intensity should drop to zero rapidly. We do not observe this behavior until the added aluminum is about 10 times that of the copper. We cannot be certain, however, with polycrystalline materials that all the added aluminum goes into solution. Furthermore, we have not considered the possible role of zinc vacancies which may behave as acceptors. Thus an excess of Al over Cu would be needed to compensate the copper acceptors.

5. SUMMARY

The infrared emission bands peaking at 0·74 and 0·80 eV in ZnS phosphors are due to the presence of Cu in the phosphors. Phosphors showing the most intense infrared emission when excited by infrared alone at −196°C (without simultaneous or previous ultraviolet excitation) can be prepared if in addition to the presence of Cu: (i) the impurity donor concentration is kept much lower than the Cu concentration; (ii) the phosphor is equilibrated at high temperatures under high sulfur pressures and then quenched rapidly to room temperature.

An interesting observation made during the course of these experiments and not reported previously is that in ZnS phosphors containing copper as the only added impurity it is possible to get a green emission band when these phosphors are equilibrated under high sulfur pressures. The orange-red emission band is always observed.[7] This effect is being investigated in more detail at the present time.

Acknowledgements—We are grateful to Mrs. D. J. Weil for preparation of the phosphors, to Dr. R. E. Halsted for measurements of spectra on many of these phosphors, and to Mr. F. C. Mostek and Dr. D. T. F. Marple for reflectance measurements reported in this paper.

REFERENCES

1. Garlick G. F. J. and Dumbleton M. J., *Proc. Phys. Soc.* B**67**, 442 (1954).
2. Browne P. F., *J. Electron.* **2**, 1 (1956–57).
3. Meijer M., *J. Phys. Chem. Solids* **7**, 153 (1958).
4. Halsted R. E., Apple E. F. and Prener J. S., *Phys. Rev. Letters* **1**, 134 (1958).
5. Pogorelyi A. D., *J. Phys. Chem., Moscow* **22**, 731 (1948).
6. Kastner J., Potter R. M. and Aven M. H., *Electrochem. Soc. Enld. Abstr.* p. 48. Electronics Division, Spring Meeting, New York, April 27–May 1, 1958.
7. Aven M. H. and Potter R. M., *J. Electrochem. Soc.* **105**, 137 (1958).
8. Broser I., *Proc. Symp. on the Role of Solid State Phenomena in Electric Circuits* p. 314. Interscience Publishers, New York (1957).
9. Klasens H. A., *J. Electrochem. Soc.* **100**, 72 (1953).
10. Apple E. F. and Williams F. E., *J. Electrochem. Soc.* **106**, 224 (1959).
11. Ullman F. G. and Dropkin J. J., *Electrochem. Soc. Enld. Abst.* p. 66. Electronics Division, Spring Meeting, April 27–May 1, 1958.
12. Potter R. M. and Aven M. H., *Bull. Amer. Phys. Soc.* (Ser. 2) **4**, 227 (1959).
13. Johnson P. D., *J. Opt. Soc. Amer.* **42**, 978 (1952).
14. Kröger F. A. and Vink H. J., *Solid State Physics* (Edited by F. Seitz and D. Turnbull) Vol. 3, p. 352. Academic Press, New York (1956).
15. Hoogenstraaten W., *J. Electrochem. Soc.* **100**, 356, (1953).

VOL. 13 1960

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 13. pp. 88–104. Printed in Great Britain.

ABSORPTION AND EMISSION SPECTRA OF AN ELECTRON IN A ONE-DIMENSIONAL DEEP TRAP

B. S. GOURARY*

Applied Physics Laboratory, The Johns Hopkins University, Silver Spring, Maryland

and

A. A. MARADUDIN†

Institute for Fluid Dynamics and Applied Mathematics, University of Maryland, College Park, Maryland

(*Received* 22 *July* 1959)

Abstract—The wave functions and the energies of the bound states of an electron in a *special* one-dimensional deep trap are calculated. Only two *essential* approximations are made in the course of the treatment, namely the Born–Oppenheimer approximation for the separation of the electronic and the nuclear motions, and the harmonic approximation for the lattice energy. Further approximations are also made, but these are not essential to the calculation and they can be avoided, if necessary. Green's function techniques are employed in the solution of the electronic and the lattice-vibration problems, and they make possible the treatment of both the localized and the extended lattice-vibration modes and their influence on the electronic absorption (emission) spectrum. The first two moments of the absorption and emission spectra are then calculated. The principal utility of this work lies in providing a test case for more approximate theories, although a generalization of the method to three dimensions and to more realistic potentials should also be possible.

1. INTRODUCTION

THE problem of the calculation of the absorption and emission spectra of an electron in a deep trap in a crystal is of considerable importance in the field of color centers, as well as in the study of the spectra of impurities in large-gap semiconductors.[1–6] A complete solution of this problem in the Born–Oppenheimer approximation requires the calculation of the wave functions of the trapped electron as functions of the nuclear coördinates and then the solution of the nuclear-motion problem, using the sum of the electronic energy and the interatomic potentials as the potential energy for the nuclear motion. Once these wave functions and energies have been evaluated, it is necessary to evaluate the probability of transition of the electron from a given electronic state and nuclear configuration to an excited electronic state and a different nuclear configuration, to sum this probability for all final states entailing the same transition energy and to average over all possible initial nuclear configurations, using a Boltzmann weighting factor. A number of difficulties stand in the way of a complete solution of this problem, and we shall now proceed to a detailed description of some of these problems.

One difficulty is presented by the determination of the electronic wave function in the Born–Oppenheimer approximation without making further unjustifiable assumptions. This is a particularly severe problem in the case of the deep trap, where the determination of the electronic wave function is usually very difficult even if the surrounding ions are assumed to be at rest in the sites which they would occupy in the perfect lattice. The requirement that the wave function be determined as a function of the lattice coördinates increases the difficulty of the problem.

Another problem is presented by the accurate determination of the vibration modes of the lattice

* Work supported by the Bureau of Ordnance, Department of the Navy, under NOrd 7386.

† This research was supported by the Air Research and Development Command of the U.S. Air Force.

VOL. 13 1960

and of their frequencies. The lattice vibration spectrum is complicated even for the perfect crystal, and its determination for the imperfect crystal poses a formidable task. The difficulty is that the lattice vibration modes are modified by the presence of a defect. Most of the modes are modified only in the immediate neighborhood of the lattice defect, but a few suffer major alterations, and the vibrations become localized around the defect. The complete solution of the deep-trap problem requires, of course, the treatment of the interaction of the electronic transition with the lattice modes, and this is impossible to carry out accurately unless the nature of the normal modes is known in the vicinity of the vacancy.

Two general approaches have been used in the past to treat this problem. One has been the configuration coördinate approach,[6] which involved the treatment of the interaction of the electronic transition with the localized modes only. It thus entailed the assumption that the influence of the extended modes on the absorption and the emission spectrum must be negligible. The usual argument presented in support of this view goes as follows. Consider an extended mode. It is easy to prove by the use of the orthonormality and completeness relations that the amplitude of an extended mode in the vicinity of the defect is of order $N^{-1/2}$, where N is the total number of particles in the crystal. Thus this interaction must vanish in the limit of an infinite crystal. The fallacy is, of course, that there are N such modes (minus a *finite* number of localized modes) and that the sum of these infinitesimal interactions does not necessarily vanish. Thus the use of a configuration coördinate alone is not adequately justified.

The other approach has been the use of the elastic-continuum representation for the crystal.[1,2,6] The electrical properties of the crystal are then taken into account by the introduction of a dielectric constant. The vibrational modes of such an elastic-dielectric continuum are then calculated, and the wave function of the electron is found approximately. The possibility of localized modes is not considered in this approximation, although it should be entirely possible to include them in such a theory. Here the fallacy is, of course, that the localized vibration modes are omitted. It is a difficulty which can be corrected, however, within the framework of the elastic dielectric continuum. The really fatal weakness of this approach is that the continuum model makes it impossible to obtain an adequate wave function for the electron. It should be possible, however, to remedy this by using the "semi-continuum" theory[6] for the electronic wave function and then considering the localized modes which exist in the presence of the Jost cavity. Even then, however, the correct shape of the vibrational spectrum cannot be obtained from the continuum theory.

Lax[3] and others[4,5] have studied the shape of the absorption and emission spectra of an electron in a general trap by the method of moments. Their work does not lessen the need for a detailed study of models, capable of yielding the parameters which enter into the expressions for the moments.

The purpose of the present work is to study a simple, completely tractable model, and to compute the first few moments of the spectrum, making only the Born–Oppenheimer approximation and the harmonic approximation for the lattice energy. This will provide us with some insight into the various factors which determine the shape of the spectrum, since there will be no involved calculational approximations to lead us astray. While the model chosen will be a highly idealized one-dimensional defect, many of the conclusions will apply to a more general class of problems. Moreover, with the use of our method it should prove possible to carry out a numerical calculation of the relative importance of the extended and the localized vibration modes in determining the shape of the spectra for this model. Such a simple but accurate calculation would provide an excellent check on more general but approximate methods.

2. DESCRIPTION OF THE MODEL AND OUTLINE OF THE CALCULATION

We study the motion of an electron in a one-dimensional potential which can be represented as a sum of localized potentials due to nuclei spaced arbitrarily along a line (see Fig. 1a), and we ask first for the electronic wave function and the energy as a function of the nuclear coördinates. Adding this energy to the potential energy of the lattice, we then expand the sum in powers of the displacements of the nuclei from their equilibrium positions, which are assumed to lie at the lattice points of a periodic one-dimensional chain, except for the

central atom, which is absent from the lattice. If we carry out this expansion only to quadratic terms in the nuclear displacements, we can then solve for the nuclear wave functions and hence

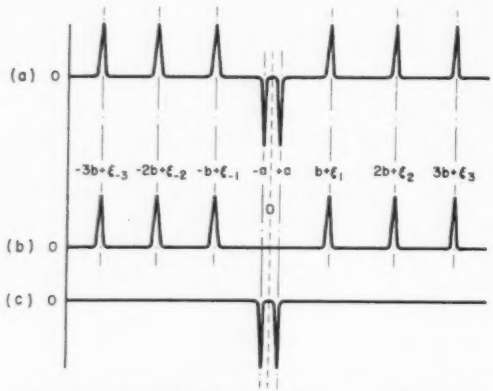


FIG. 1. (a) The potential energy of the trapped electron, that is $V(x)+U(x)$. (b) $V(x)$. (c) $U(x)$.

obtain the total wave function for the electron-lattice system in the Born–Oppenheimer approximation.

Since in solving the electronic part of this problem we will be primarily interested in the wave functions and energy levels of an electron bound to the vacancy, we find it most convenient to formulate the electronic problem in terms of the Green's function solution to Schrödinger's equation for the motion of an electron in the one-dimensional, aperiodic potential mentioned above. It is our purpose in this section to present a general outline of the manner in which such a Green's function can be constructed, and to derive the expressions for the electronic wave functions and energies in terms of this function. The use of these methods with a particular choice for the nuclear potential will be described in the following section. The calculation of the vibration modes of the lattice in the presence of the vacancy is also most conveniently carried out with the aid of Green's function techniques, and a general discussion of such methods, together with their application to the present problem, will be given in Section 4.

The Schrödinger equation for an electron moving in an aperiodic potential due to nuclei spaced arbitrarily along a line and a vacancy at the origin (see Fig. 1a) is, in Hartree atomic units

$$\frac{1}{2}\left(\frac{d^2}{dx^2}-\alpha^2-2V(x)\right)\psi(x)=U(x)\psi(x). \quad (2.1)$$

where

$$-\frac{\alpha^2}{2}=E \quad (2.2)$$

is the energy of the electron, $V(x)$ represents the potential due to the nuclei (Fig. 1b) and $U(x)$ is the potential due to the vacancy and will be assumed to be localized (see Fig. 1c). In the present problem we will be primarily interested in those solutions of equation (2.1) which correspond to an electron localized in the neighborhood of the defect. The solution to equation (2.1) is most readily carried out with the aid of the Green's function for the differential operator on the left-hand side of this equation. This Green's function satisfies the equation:

$$\frac{1}{2}\left(\frac{d^2}{dx^2}-\alpha^2-2V(x)\right)G(x,x')=\delta(x-x'). \quad (2.3)$$

This equation cannot be solved in closed form for a general $V(x)$, but can be expressed in a form convenient for our purposes in terms of the free-particle Green's function $g(x, x')$, which satisfies

$$\frac{1}{2}\left(\frac{d^2}{dx^2}-\alpha^2\right)g(x,x')=\delta(x-x'). \quad (2.4)$$

Assuming the following form for $g(x, x')$,

$$g(x,x')=\frac{1}{2\pi}\int_{-\infty}^{\infty}A(y)\exp[i(x-x')y]\,dy. \quad (2.5)$$

we find on substituting this expression into equation (2.4) that

$$A(y)=\frac{-2}{y^2+\alpha^2}$$

so that

$$g(x,x')=-\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\exp[i(x-x')y]}{y^2+\alpha^2}\,dy$$

$$=-\frac{1}{\alpha}\exp(-\alpha|x-x'|)=g(|x-x'|). \quad (2.6)$$

VOL. 13 1960

The relation between $G(x, x')$ and $g(x, x')$ is obtained by rewriting equation (2.3) as

$$\frac{1}{2}\left(\frac{d^2}{dx^2}-\alpha^2\right)G(x, x') = V(x)G(x, x')+\delta(x-x') \tag{2.7}$$

and using equation (2.4) to cast this in the form of an integral equation

$$G(x, x') = g(x, x')+\int_{-\infty}^{\infty} g(x, x_1)V(x_1)G(x_1, x')\,dx_1. \tag{2.8}$$

This integral equation for $G(x, x')$ can be solved by iteration with the result that

$$G(x, x') = g(x, x')+\int_{-\infty}^{\infty} g(x, x_1)V(x_1)g(x_1, x')\,dx_1+$$

$$+\int_{-\infty}^{\infty} dx_1 \int_{-\infty}^{\infty} dx_2 g(x, x_1)V(x_1)g(x_1, x_2)V(x_2)g(x_2x')+ \ldots \tag{2.9}$$

In the case that the potential $V(x)$ can be written as a sum of contributions from many scattering centers,

$$V(x) = \sum_n \phi_n(x-x_n), \tag{2.10}$$

equation (2.9) can be written as

$$G(x, x') = g(x, x')+\sum_n \int_{-\infty}^{\infty} g(x, x_1)\phi_n(x_1-x_n)g(x_1, x')\,dx_1+$$

$$+\sum_{m,n} \int_{-\infty}^{\infty} dx_1 \int_{-\infty}^{\infty} dx_2 g(x, x_1)\phi_m(x_1-x_m)g(x_1, x_2)\phi_n(x_2-x_n)g(x_2, x')+ \ldots \tag{2.11}$$

An evaluation of this expression for a special choice of $\phi_n(x)$ will be carried out in the next section.

We now return to an investigation of equation (2.1). Casting it into the form of an integral equation, we find

$$\psi(x) = \int_{-\infty}^{\infty} G(x, x')U(x')\psi(x')\,dx'. \tag{2.12}$$

In the special case of interest to us, we choose

$$U(x) = -U_0[\delta(x+a)+\delta(x-a)]. \tag{2.13}$$

The delta function is chosen in order to convert the integral equation into a set of algebraic equations. The need for *two* delta functions arises from the desire to have *two* bound states. The integral equation (2.12) reduces to

$$\psi(x) = -U_0[G(x, -a)\psi(-a)+G(x, a)\psi(a)], \tag{2.14}$$

and the consistency equations become

$$\begin{pmatrix} U_0G(-a, -a)+1 & U_0G(-a, a) \\ U_0G(a, -a) & U_0G(a, a)+1 \end{pmatrix} \times \begin{pmatrix} \psi(-a) \\ \psi(a) \end{pmatrix} = 0. \tag{2.15}$$

The condition that this homogeneous system of equations have a solution for $\psi(a)$ and $\psi(-a)$ is the vanishing of the determinant of their coefficients. This condition leads to the following equation:

$$\frac{1}{U_0} = -\tfrac{1}{2}[G(a, a)+G(-a, -a)]\pm$$

$$\pm\tfrac{1}{2}\sqrt{\{[G(a, a)-G(-a, -a)]^2+4G^2(a, -a)\}}, \tag{2.16}$$

which relates the energy parameter α to the nuclear

coördinates. The solution of these equations will be carried out for a special case in the next section.

3. THE ELECTRONIC PROBLEM

We are now in a position to apply the formalism developed in the preceding section to a concrete example. The potential $V(x)$, which represents the unperturbed potential for the motion of the electron in our lattice, is chosen to be (see Fig. 1b)

$$V(x) = v \sum_{n=-\infty}^{\infty}{}' \delta(x-x_n), \tag{3.1}$$

where the prime on the summation excludes the value $n = 0$, and x_n is the coördinate of the nth nucleus.* The Green's function $G(x, x')$ in this case becomes

$$G(x, x') = g(x, x') + v \sum_{n}{}' g(x, x_n)g(x_n, x') + v^2 \sum_{m,n}{}'' g(x, x_m)g(x_m, x_n)g(x_n, x') + \ldots . \tag{3.2}$$

Substituting the result of equation (2.6) into this equation, we obtain

$$\begin{aligned} G(x, x') = &-\frac{1}{\alpha}\exp(-\alpha|x-x'|) + \frac{v}{\alpha^2}\sum_{n}{}'\exp(-\alpha|x-x_n|-\alpha|x_n-x'|) - \\ &-\frac{v^2}{\alpha^3}\Big[\sum_{n}{}'\exp(-\alpha|x-x_n|-\alpha|x_n-x'|) + \sum_{\substack{m,n\\ m\neq n}}{}''\exp(-\alpha|x-x_m|-\alpha|x_m-x_n|-\alpha|x_n-x'|)\Big] + \\ &+\frac{v^3}{\alpha^4}\Big[\sum_{n}{}'\exp(-\alpha|x-x_n|-\alpha|x_n-x'|) + 2\sum_{\substack{m,n\\ m\neq n}}{}''\exp(-\alpha|x-x_m|-\alpha|x_m-x_n|-\alpha|x_n-x'|) + \\ &+\sum_{\substack{m,n\\ m\neq n\neq p}}{}''\exp(-\alpha|x-x_m|-\alpha|x_m-x_n|-\alpha|x_n-x_p|-\alpha|x_p-x'|)\Big] + \ldots \end{aligned} \tag{3.3}$$

Combining all terms with the same summand, we obtain the following exact result:

$$\begin{aligned} G(x, x') = &-\frac{1}{\alpha}\Big[\exp(-\alpha|x'-x|) - \frac{(v/\alpha)}{1+(v/\alpha)}\sum_{n}{}'\exp(-\alpha|x'-x_n|-\alpha|x_n-x|) + \\ &+\frac{(v/\alpha)^2}{(1+v/\alpha)^2}\sum_{\substack{m,n\\ m\neq n}}{}''\exp(-\alpha|x'-x_m|-\alpha|x_m-x_n|-\alpha|x_n-x|) - \ldots\Big] \end{aligned} \tag{3.4}$$

The convergence of this modified iteration series will be established in a subsequent paper, which deals with impurity band conduction.

With reasonable choices for the parameters v, U_0 and the lattice spacing b, we can insure that the second sum in this expression is negligible in comparison with the first two terms. Thus in all that follows we use the following expression for $G(x, x')$:

$$G(x, x') = -\frac{1}{\alpha}\Big[\exp(-\alpha|x'-x|) - \frac{(v/\alpha)}{[1+(v/\alpha)]}\sum_{n}{}'\exp(-\alpha|x'-x_n|-\alpha|x_n-x|)\Big]. \tag{3.5}$$

Since the only values of the Green's function

* This is a modified version of a model first used by KRONIG and PENNEY. For a discussion of their work see for example, SEITZ[7].

which arise in the calculation of the bound-state energies are $G(a, a)$, $G(a, -a)$, $G(-a, a)$ and $G(-a, -a)$, where $a < b$, we need consider only the general expression for $G(x, x')$ for $|x|, |x'| < b$. We begin by putting

$$x_n = nb + \xi_n \quad (n = \pm 1, \pm 2, \pm 3, \ldots) \tag{3.6}$$

into equation (3.5), where $|\xi_n| \ll b$. With this substitution, equation (3.5) becomes

$$G(x, x') = -\frac{1}{\alpha}\Big[\exp(-\alpha|x'-x|) - \frac{(v/\alpha)}{[1+v/\alpha]}\sum_{n=1}^{\infty}\exp[-\alpha(2nb+2\xi_n-x'-x)] -$$

$$-\frac{(v/\alpha)}{[1+v/\alpha]}\sum_{n=1}^{\infty}\exp[-\alpha(2nb-2\xi_{-n}+x'+x)]\Big]. \tag{3.7}$$

We now expand the right-hand side of this equation in powers of ξ_n and ξ_{-n}, with the result that

$$G(x, x') = -\frac{1}{\alpha}\Big[\exp(-\alpha|x'-x|) - \Big(\frac{1}{[1+(\alpha/v)]}\Big)\sum_{n=1}^{\infty}\exp[-\alpha(2nb-x'-x)](1-2\alpha\xi_n +$$

$$+2\alpha^2\xi_n^2 - \ldots) - \Big(\frac{1}{1+(\alpha/v)}\Big)\sum_{n=1}^{\infty}\exp[-\alpha(2nb+x'+x)](1+2\alpha\xi_{-n} +$$

$$+2\alpha^2\xi_{-n}^2 + \ldots)\Big] = -\frac{1}{\alpha}\Big[\exp(-\alpha|x'-x|) - \frac{2\cosh\alpha(x+x')}{1+(\alpha/v)}\frac{1}{\exp(2\alpha b)-1}\Big] +$$

$$+\frac{1}{\alpha}\frac{1}{1+(\alpha/v)}\sum_{n=1}^{\infty}\exp(-2n\alpha b)\{\exp[\alpha(x+x')](-2\alpha\xi_n+2\alpha^2\xi_n^2) + \exp[-\alpha(x+x')](2\alpha\xi_{-n}+2\alpha^2\xi_{-n}^2)\} \tag{3.8}$$

Since the dependence of the electronic energy on the nuclear coördinates will be slight compared with the sensitivity of the lattice energy to the nuclear positions, its principal effect will be to shift the equilibrium positions of the nuclei away from $x_n = nb$. The potential energy of the perfect lattice contains no linear terms in the ξ_n, due to the equilibrium condition, so that the entire static distortion of the lattice will arise from the linear terms in the ξ_n's in the expansion of the electronic energy. The terms quadratic in the ξ_n's will combine with corresponding terms in the expansion of the lattice potential energy to alter the normal mode frequencies of the lattice. Since the electronic contribution will be smaller than the lattice term, we neglect it for convenience. This neglect is not essential for the calculation, nor does it influence any of our major conclusions. It does have the simplifying feature of causing the normal mode frequencies of the lattice to be the same whether the electron is in its ground state or in its first excited state. This approximation means that

$$G(x, x') \approx -\frac{1}{\alpha}\Big[\exp(-\alpha|x'-x|) - \frac{2}{1+\alpha/v}\frac{\cosh\alpha(x+x')}{\exp(2\alpha b)-1}\Big] -$$

$$-\frac{2}{1+\alpha/v}\sum_{n=1}^{\infty}\exp(-2n\alpha b)\{\exp[\alpha(x+x')]\xi_n - \exp[-\alpha(x+x')]\xi_{-n}\} \tag{3.9}$$

VOL. 13 1960

The equations (2.16), expressing the electron energy implicitly in terms of the nuclear displacements, now become

$$\frac{1}{U_0}=\frac{1}{\alpha}[1\pm\exp(-2\alpha a)]-$$

$$-\frac{2}{\alpha}\frac{(\cosh 2\alpha a\pm 1)}{(1+\alpha/v)\,[\exp(2\alpha b)-1]}+$$

$$+\frac{2(\cosh 2\alpha a\pm 1)}{(1+\alpha/v)}\sum_{n=1}^{\infty}\exp(-2n\alpha b)(\xi_n-\xi_{-n}) \tag{3.10}$$

to first order in the nuclear displacements. Here the upper sign is associated with the ground state and the lower sign with the excited state. This can be seen from the following argument. The consistency equations (2.15) for an unstrained ($\xi_n=0$) lattice can be written as

$$\begin{pmatrix}A & B\\ B & A\end{pmatrix}\begin{pmatrix}\psi(a)\\ \psi(-a)\end{pmatrix}=0, \tag{3.11}$$

where

$$A=\frac{1}{U_0}+G(a,a)\qquad B=G(-a,a)=G(a,-a)$$

$$=\frac{1}{U_0}+G(-a,-a) \tag{3.12}$$

This matrix is diagonalized by the unitary matrix

$$S=\frac{1}{\sqrt{(2)}}\begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix} \tag{3.13}$$

Now,

$$\frac{1}{2}\begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix}\begin{pmatrix}A & B\\ B & A\end{pmatrix}\begin{pmatrix}1 & 1\\ -1 & 1\end{pmatrix}\left[\frac{1}{\sqrt{(2)}}\begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix}\begin{pmatrix}\psi(a)\\ \psi(-a)\end{pmatrix}\right]$$

$$=\begin{pmatrix}A-B & 0\\ 0 & A+B\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{(2)}}[\psi(a)-\psi(-a)]\\ \frac{1}{\sqrt{(2)}}[\psi(a)+\psi(-a)]\end{pmatrix}. \tag{3.14}$$

Numerical solutions show that the state defined by $A+B=0$ has the lower energy, while the state defined by $A-B=0$ has the higher energy. The wave function in the former case is clearly even, while it is odd in the latter case. In the case of finite ξ_n, the symmetries are no longer simply odd or even, but the relative magnitudes of the energies will be preserved.

For the further development of this problem we require not so much the electronic energy with its implicit dependence on the ξ_n's but its explicit dependence in the form

$$E=E_0+\sum_{n=1}^{\infty}(E_n\xi_n+E_{-n}\xi_{-n})+\ \ldots \tag{3.15}$$

where, in keeping with our earlier remarks, we have retained only terms linear in the ξ_n's.

This inversion problem is carried out in Appendix A, and the results for the ground-state and excited-state energies are

$$E_g=E_0^{(g)}+\frac{\alpha^{(g)}}{f_0'(\alpha^{(g)})}\frac{2(\cosh 2a\alpha^{(g)}+1)}{1+\alpha^{(g)}/v}\sum_{n=1}^{\infty}\exp(-2n\alpha^{(g)}b)(\xi_n-\xi_{-n}). \tag{3.16a}$$

$$E_e=E_0^{(e)}+\frac{\alpha^{(e)}}{f_0'(\alpha^{(e)})}\frac{2(\cosh 2a\alpha^{(e)}-1)}{1+\alpha^{(e)}/v}\sum_{n=1}^{\infty}\exp(-2n\alpha^{(e)}b)(\xi_n-\xi_{-n}). \tag{3.16b}$$

where $E_0^{(g)}$ and $E_0^{(e)}$ are obtained from the solutions $\alpha^{(g)}$, $\alpha^{(e)}$, to the zero-order equations

$$\frac{1}{U_0}=\frac{1}{\alpha^{(g)}}[1+\exp(-2\alpha^{(g)}a)]-\frac{2}{\alpha^{(g)}}\frac{(\cosh 2\alpha^{(g)}a+1)}{\left(1+\frac{\alpha^{(g)}}{v}\right)[\exp(2\alpha^{(g)}b)-1]}. \tag{3.17a}$$

$$\frac{1}{U_0}=\frac{1}{\alpha^{(e)}}[1-\exp(-2\alpha^{(e)}a)]-\frac{2}{\alpha^{(e)}}\frac{(\cosh 2\alpha^{(e)}a-1)}{\left(1+\frac{\alpha^{(e)}}{v}\right)[\exp(2\alpha^{(e)}b)-1]}, \tag{3.17b}$$

respectively, through the relation

$$E=-\tfrac{1}{2}\alpha^2. \tag{3.18}$$

$f_0'(\alpha)$ is a constant which is defined in Appendix A.

From equations (3.16), we see immediately that the force exerted on the nth atom, F_n, is equal and opposite to the force exerted on the $(-n)$th atom:

$$F_n=-\left(\frac{\partial E}{\partial \xi_n}\right)_{\xi=0}=-F_{-n}=\left(\frac{\partial E}{\partial \xi_{-n}}\right)_{\xi=0}, \tag{3.19}$$

and it decays exponentially with increasing n.

The wave functions for the electron in its ground and excited states are given by

$$\psi_g(x)=-U_0\psi_g(a)[G_g(x,-a)+G_g(x,a)]. \tag{3.20a}$$

$$\psi_e(x)=-U_0\psi_e(a)[-G_e(x,-a)+G_e(x,a)]. \tag{3.20b}$$

where $G_g(x,x')$ and $G_e(x,x')$ are obtained from equation (3.8) on replacing α by $\alpha^{(g)}$ and $\alpha^{(e)}$, respectively.

4. LATTICE VIBRATIONS IN THE PRESENCE OF A DEFECT

For the purposes of the present work, it will be necessary for us to choose a particular model of the lattice which, on the one hand, will be completely tractable even in the presence of an imperfection, while, on the other hand, it will lead to a finite result for the dispersion of an atom about its equilibrium point, since the latter quantity will appear in certain moments of the absorption spectrum. For these reasons, we shall consider the following model, in which only longitudinal vibrations are allowed. The spacing between successive atoms in the one-dimensional crystal is b. Each atom is tied to each of its two nearest neighbors by a spring of spring constant k_2. It is also tied to its lattice point by a spring of spring constant k_1. This latter coupling gives the model some three-dimensional character. The equation of motion of the mth atom

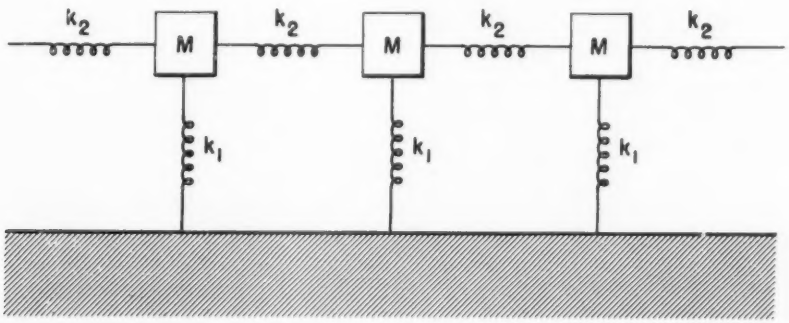


FIG. 2. The model employed in the calculation of the lattice vibrations. It should be noted that only longitudinal vibrations are allowed in this model.

in the perfect lattice is therefore

$$M\ddot{u}_m = -k_1u_m - k_2(2u_m - u_{m-1} - u_{m+1}), \quad (4.1)$$

where M is the mass of the atom. These equations can be readily solved if periodic boundary conditions are introduced, i.e. if the following condition is imposed

$$u_{m+N} = u_m, \quad (4.2)$$

where N is the number of atoms in the chain. The solutions can be obtained in various forms. For our purposes, it will be convenient to use standing waves, i.e. solutions of the form

$$u_m^\lambda = A\begin{Bmatrix}\sin(m\theta)\\ \cos(m\theta)\end{Bmatrix}\cos(\omega_\lambda t - \phi) = w_m^\lambda \cos(\omega_\lambda t - \phi) \quad (4.3)$$

where A is a constant, ω_λ is the angular frequency of the λth mode, ϕ is a constant and θ is given by the boundary conditions as

$$\theta = 2\pi p_\lambda / N, \quad (4.4)$$

p_λ can take on the values $p_\lambda = 0, 1, 2, \ldots N/2$ for the cos $m\theta$ solution and $p_\lambda = 1, 2, 3 \ldots N/2-1$ for the sin $m\theta$ solution, where N has been assumed to be even for convenience. Thus there is a total of N modes. The normal mode frequencies can be obtained by substituting equation (4.3) into (4.1). Thus

$$\omega = [M^{-1}(k_1 + 2k_2 - 2k_2\cos\theta)]^{1/2}. \quad (4.5)$$

Clearly the lowest frequency occurs for $\theta = 0$, and the highest frequency arises for $\theta = \pi$. The spectrum of ω^2 has the same shape as in the more familiar case where $k_1 = 0$, but it is displaced by k_1/M upward from the origin. Thus zero is not a natural frequency of the system. This fact will turn out to have considerable practical importance.

In order to simplify the treatment of the normal modes in the presence of an imperfection, it will be convenient to introduce the operator.

$$L_{mn} = (k_1 + 2k_2)\delta_{mn} - k_2(\delta_{m-1,n} + \delta_{m+1,n}) = L_{nm}. \quad (4.6)$$

Then the equations of motion of the atom in the perfect lattice can be written in the form

$$M\ddot{u}_m + \sum_n L_{mn}u_n = 0. \quad (4.7)$$

The equations of motion of the atom in the presence of an imperfection are

$$M\ddot{u}_m + \sum_n L_{mn}u_n = F_m \quad (4.8)$$

where F_m is the force due to the defect acting on the mth atom. We shall require the complete solution of equation (4.8) i.e. both the vibrational frequency and the eigenvector of each mode.

Following MONTROLL and POTTS[8], we shall solve the equations (4.8) by a Green's function method. The Green's function $\gamma_{mn}(\omega)$ will satisfy the equations

$$-M\omega^2\gamma_{mk}(\omega) + \sum_n L_{mn}\gamma_{nk}(\omega) = -\delta_{mk}. \quad (4.9)$$

Then the set u_m defined by

$$u_n = w_n \cos(\omega t - \phi) \quad (4.10)$$

$$w_n = -\sum_k \gamma_{nk}(\omega)F_k \quad (4.11)$$

will be a solution of equation (4.8).

The explicit form of $\gamma_{nk}(\omega)$ is derived in Appendix B, and several expressions for it are given there. It is a function of ω and of $|n-k|$ only, and can therefore be written as $\gamma_{n-k}(\omega)$.

In the presence of a lattice imperfection, the forces F_n are linear in the displacements, w_n, of a few atoms surrounding the imperfection. They can be written in the form

$$F_k = \sum_t C_{kt}w_t, \quad (4.12)$$

where the C_{kt} are the expansion coefficients. The expression (4.11) then becomes

$$w_n = -\sum_t \left[\sum_k \gamma_{nk}(\omega)C_{kt}\right]w_t, \quad (4.13)$$

which is a set of linear equations relating *all* the w_n's to a *few* w_t's near the imperfection. For example, if the mass of only one atom is changed, then the indices k and t assume only one value. If a force constant between two atoms is changed, the indices k and t assume the two values corresponding to the two atoms involved. If n is now permitted to assume the small number of values which form the range of t, then a set of simultaneous linear equations (the consistency equations) is obtained for the displacements of the several atoms neighboring the imperfection. These consistency equations have a solution only if the determinant of

the coefficients of the system vanishes, and the roots of the determinant give us the natural frequencies of the system, which we call ω_λ. The N eigenvectors w_n obtained by substituting the ω_λ's and the corresponding w_l's into equation (4.13) can be normalized to unity. They will then obey the orthogonality and the completeness relations

$$\sum_\lambda w_n^\lambda w_m^\lambda = \delta_{nm} \tag{4.14}$$

$$\sum_n w_n^\lambda w_n^{\lambda'} = \delta_{\lambda\lambda'}. \tag{4.15}$$

In the present paper, we shall be interested in the normal modes which arise in the presence of a simple defect, consisting of the removal of one atom and the insertion of a force constant between its two nearest neighbors. Consider the case where the atom at 0 is removed and its interactions with its lattice point and with its nearest neighbors are cancelled. Let us then insert a new force constant between the atom at 1 and the atom at -1, and let us call this force constant K. Then,

$$F_0 = -M\omega^2 w_0 + (k_1+2k_2)w_0 - k_2(w_1+w_{-1}) \tag{4.16}$$

$$F_1 = k_2(w_1-w_0) - K(w_1-w_{-1}) \tag{4.17}$$

$$F_{-1} = k_2(w_{-1}-w_0) - K(w_{-1}-w_1). \tag{4.18}$$

This leads to the following expression for w_n:

$$w_n = \delta_{n0}w_0 - w_1[k_2(\gamma_{n-1}(\omega)-\gamma_n(\omega)) + K(\gamma_{n+1}(\omega)-\gamma_{n-1}(\omega))] - \\ - w_{-1}[k_2(\gamma_{n+1}(\omega)-\gamma_n(\omega)) + K(\gamma_{n-1}(\omega)-\gamma_{n+1}(\omega))]. \tag{4.19}$$

Clearly w_0 drops out. We can thus get our set of simultaneous linear equations by setting $n = 1$ or -1. It is easy to simplify this problem by noticing that the resulting normal modes are likely to be either symmetric or antisymmetric about the defect. This suggests that the matrix of the coefficients may be diagonalized by the transformation matrix

$$S = \frac{1}{\sqrt{(2)}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \tag{4.20}$$

The matrix of the coefficients is

$$\begin{pmatrix} A & B \\ B & A \end{pmatrix}. \tag{4.21}$$

where

$$A = 1 + k_2[\gamma_0(\omega)-\gamma_1(\omega)] + K[\gamma_2(\omega)-\gamma_0(\omega)] \tag{4.22}$$

$$B = k_2[\gamma_2(\omega)-\gamma_1(\omega)] + K[\gamma_0(\omega)-\gamma_2(\omega)]. \tag{4.23}$$

The transformed equations are then

$$\begin{pmatrix} A-B & 0 \\ 0 & A+B \end{pmatrix}\begin{pmatrix} w_1-w_{-1} \\ w_1+w_{-1} \end{pmatrix} = 0. \tag{4.24}$$

There are then two cases. Either

$$(1)\ A+B = 0,\ A-B \neq 0, \text{ and } w_1 = w_{-1}. \tag{4.25}$$

or

$$(2)\ A-B = 0,\ A+B \neq 0, \text{ and } w_1 = -w_{-1}. \tag{4.26}$$

It is easy to show that case (1) does not lead to any localized solutions. To see this, write out the condition obtained by assuming that $w_1 = w_{-1}$ and that $n = 0$ in equation (4.19) and use equations (B.7) and (B.8). The interesting case is (2). Here, equation (4.19) becomes

$$w_n = w_1(2K-k_2)[\gamma_{n-1}(\omega)-\gamma_{n+1}(\omega)] = -w_{-n}. \tag{4.27}$$

and the eigenvalue equation can be written in the form

$$\gamma_0(\omega)-\gamma_2(\omega) = (2K-k_2)^{-1}. \tag{4.28}$$

This gives rise to one localized mode, and to a frequency shift of the allowed modes in the band.

The description of the emission and absorption spectra which we shall present in Section 6 will consist of the calculation of the first three moments of the spectra. This will require the sum of the series

$$\sum_\lambda w_m^\lambda w_n^\lambda \omega_\lambda^{-2}, \tag{4.29}$$

which is essentially the bilinear expansion of a Green's function $\Gamma_{mn}(0)$. This function $\Gamma_{mn}(\omega)$ is

VOL. 13 1960

defined as the Green's function for the perturbed lattice; that is, it satisfies the equation

$$-M\omega^2\Gamma_{mn}(\omega)+\sum_k L_{mk}\Gamma_{kn}(\omega)-\\ -\sum_k C_{mk}\Gamma_{kn}(\omega)=-\delta_{mn} \qquad (4.30)$$

where the C_{mk}'s describe the particular perturbation. We shall first obtain the bilinear expansion for this Green's function, and then we shall find a closed form for the same function. Equating these two expressions, we shall obtain a closed form expression for the series.

The bilinear expansion for the Green's function is obtained as follows. We assume that the Green's function has an expansion of the form

$$\Gamma_{kn}(\omega)=\sum_\lambda a_\lambda w_k^\lambda w_n^\lambda; \quad a_\lambda=a_\lambda(\omega). \qquad (4.31)$$

where the a_λ's are undetermined coefficients, and obtain expressions for the a_λ's by requiring that equation (4.31) should satisfy (4.30). Using equations (4.14) and (4.15) and the eigenvalue equation for the w_n's, namely

$$\sum_k [L_{mk}-C_{mk}]w_k^\lambda = M\omega_\lambda^2 w_m^\lambda \qquad (4.32)$$

we obtain

$$a_\lambda(\omega)=[M(\omega^2-\omega_\lambda^2)]^{-1}. \qquad (4.33)$$

Thus clearly the expression (4.29) is simply

$$-M\Gamma_{mn}(0). \qquad (4.34)$$

Our next task is to obtain a closed form for equation (4.34). This can be done in the following manner. We multiply equation (4.30) by $\gamma_{km}(\omega)$ and sum over m; we then multiply equation (4.9) by $\Gamma_{nm}(\omega)$ and sum over m. Subtracting the second equation from the first, we obtain

$$\Gamma_{kn}(\omega)=\gamma_{kn}(\omega)-\sum_m\sum_t \gamma_{km}(\omega)C_{mt}\Gamma_{tn}(\omega). \qquad (4.35)$$

We shall be concerned only with localized defects, of course, and consequently the summation over t goes over only a few values while n may take on any values. Setting k equal to these values gives us a set of simultaneous linear equations for $\Gamma_{tn}(\omega,)$ where t is restricted to the few atoms constituting the defect, while n may take on any value. This set of equations has the same determinant of coefficients as the set resulting from equation (4.13), but since it is inhomogeneous, it will have solutions for those values of ω for which the determinant does *not* vanish. This means that it will have solutions for all frequencies except the ω_λ's. The explicit solution is obtained by Cramer's rule. Since $\omega=0$ is not an eigenvalue of the unperturbed lattice, and since the localized mode cannot correspond to a pure translation of the lattice, it follows that this solution will exist for $\omega=0$, thus providing us with a closed form for $\Gamma_{tn}(0)$. $\Gamma_{kn}(0)$ for all values of k can then be obtained from equation (4.35). Thus we have obtained our summation formula.

For the defect described by equations (4.16), (4.17) and (4.18), we have

$$\Gamma_{nm}(\omega)+\Gamma_{1m}(\omega)[k_2(\gamma_{n-1}(\omega)-\gamma_n(\omega))+K(\gamma_{n+1}(\omega)-\gamma_{n-1}(\omega))]+\\ +\Gamma_{-1m}(\omega)[k_2(\gamma_{n+1}(\omega)-\gamma_n(\omega))+K(\gamma_{n-1}(\omega)-\gamma_{n+1}(\omega))]=\gamma_{nm}(\omega)$$

for all values of n and m, except 0 which never occurs (the particle at the origin is missing). Setting n equal to 1 and then to -1, we obtain the set of simultaneous equations for $\Gamma_{lm}(\omega)$ and $\Gamma_{lm}(\omega)$, which leads to the result

$$\begin{pmatrix}\Gamma_{1m}(\omega)\\ \Gamma_{-1m}(\omega)\end{pmatrix}=(A^2-B^2)^{-1}\begin{pmatrix}A & -B\\ -B & A\end{pmatrix}\begin{pmatrix}\gamma_{m-1}(\omega)\\ \gamma_{m+1}(\omega)\end{pmatrix}. \qquad (4.37)$$

where A and B are defined by equations (4.22) and (4.23). Thus equations (4.36) and (4.37) furnish a closed form for the Green's function for the perturbed lattice. This result thus enables us to sum the series (4.29).

5. TOTAL WAVE FUNCTION AND ENERGY OF THE SYSTEM

We are now ready to study the wave function and

VOL. 13 1960

the energy of the entire system, consisting of an electron moving in the potential of a one-dimensional vibrating lattice. The lattice interacts with the electron by means of delta function repulsive potentials. The 0th lattice atom is missing, and the resulting vacancy interacts with the electron by means of two attractive delta-function potentials. The interactions of the missing atom with its neighbors and with its lattice point are absent, of course, but there is an added interaction between the atom at $+1$ and the atom at -1. The electronic wave function as a function of the nuclear coördinates is given by equations (3.20a) and (3.20b). The electronic energy is given by equations (3.16a) and (3.16b). The total potential energy for the nuclear motion is, of course, the sum of the electronic energy and the energy of interaction among the nuclei. We shall denote this total potential energy by W, and will write its expansion in terms of the displacements of the atoms from equilibrium. Thus the energy expansion is

$$W = W^0 + \sum_{n=-\infty}^{\infty}{}' E_n\xi_n + \\ +(\tfrac{1}{2})\sum_{n,m=-\infty}^{\infty}{}'' (L_{nm}-C_{nm})\xi_n\xi_m \qquad (5.1)$$

where the prime(s) indicate the omission of the point $n, m = 0$. The correctness of this equation should be obvious from the inspection of equations (3.15) and (4.32).

We shall now expand the displacement of the nth atom ξ_n in terms of the normal coördinates of the system. Thus

$$\xi_n = M^{-1/2}\sum_\lambda Q_\lambda w_n^\lambda. \qquad (5.2)$$

Substituting this into equation (5.1) and utilizing equations (4.14), (4.15) and (4.32) we have

$$W = W^0 - \sum_\lambda Q_\lambda Q_\lambda^0\omega_\lambda^2 + (\tfrac{1}{2})\sum_\lambda Q_\lambda^2\omega_\lambda^2 \\ = W^0 - (\tfrac{1}{2})\sum_\lambda (Q_\lambda^0)^2\omega_\lambda^2 + \\ +(\tfrac{1}{2})\sum_\lambda \omega_\lambda^2(Q_\lambda - Q_\lambda^0)^2 \qquad (5.3)$$

where

$$Q_\lambda^0 = -\omega_\lambda^{-2}M^{-1/2}\sum_n{}' E_n w_n^\lambda, \qquad (5.4)$$

and it depends on the state of the electron. The kinetic energy of the lattice is likewise

$$(M/2)\sum_n \dot{\xi}_n^2 = (\tfrac{1}{2})\sum_\lambda \dot{Q}_\lambda^2 \to -(\tfrac{1}{2})\sum_\lambda \left(\frac{\partial}{\partial Q_\lambda}\right)^2. \qquad (5.5)$$

Thus the total Hamiltonian of the electron plus lattice system consists of a quantity independent of the lattice coördinates, and a part which is a sum of harmonic-oscillator-type Hamiltonians for each of the displaced normal modes $(Q_\lambda - Q_\lambda^0)$ of the perturbed lattice. The quantization of such a system is trivial, of course, and the eigenvalues of the total Hamiltonian can be written in terms of the energy eigenvalues of the individual normal modes. Thus the eigenvalues are

$$\mathscr{H}' = W^0 - (\tfrac{1}{2})\sum_\lambda \omega_\lambda^2(Q_\lambda^0)^2 + \sum_\lambda \omega_\lambda(n_\lambda + 1/2) \qquad (5.6)$$

The wave function of the λth mode in the kth state is

$$\phi_k^\lambda(Q_\lambda - Q_\lambda^0) = N_k^\lambda H_k[\sqrt{(\omega_\lambda)}(Q_\lambda - Q_\lambda^0)]\times \\ \times \exp\left[-\frac{\omega_\lambda}{2}(Q_\lambda - Q_\lambda^0)^2\right] \qquad (5.7)$$

where N_k is a normalizing factor and H_k is the kth Hermite polynomial.

6. ABSORPTION AND EMISSION SPECTRA

The next step in our calculation is the derivation of the shape of the absorption and emission spectra. Following LAX[3], we shall confine ourselves to the calculation of the first few moments of the spectra. This will allow us to obtain relatively simple expressions for the measurable quantities, namely, the zeroth, first and second moments of the spectrum.

Before we embark upon this calculation, let us state the explicit significance of W^0, which will be required in the calculation of the transition energy. This quantity, which first occurs in equation (5.1), is the sum of the electronic and lattice energies when the atoms are in the perfect-lattice positions. If we choose as our energy zero the state of the system when all the lattice coördinates are at zero and the trapped electron is removed to infinity, then W^0 is just the electronic energy, i.e.

$$W^0 = E_0^{(g)} \text{ for the ground state} \qquad (6.1)$$

and

$$W^0 = E_0^{(e)} \text{ for the excited state.} \qquad (6.2)$$

The E_0's are defined in equations (3.16a) and (3.16b).

The probability of an allowed transition is proportional to the square of the electric dipole matrix element evaluated between the initial-state wave function and the final-state wave function. Thus the normalized absorption spectrum of the system is

$$I_{ge}(\omega) = B \sum_n P(n) \sum_{n'} |(n', e|M|n, g)|^2 \delta[\mathscr{H}'(n'e) - \mathscr{H}'(ng) - \omega]. \qquad (6.3)$$

where M is the dipole moment operator, n goes over all the initial vibrational states, n' goes over all the final vibrational states, $P(n)$ is the statistical weight factor for the initial vibrational states and B is the normalizing factor.

Following LAX[3], we shall write the characteristic function of equation (6.3), and then expand it to obtain the moments of the spectrum. Thus,

$$\begin{aligned} C_{ge}(t) &= \int_{-\infty}^{\infty} I_{ge}(\omega) \exp(it\omega)\, d\omega = \sum_{r=0}^{\infty} \frac{(it)^r}{r!} \langle \omega^r \rangle \\ &= B \sum_n P(n) \sum_{n'} (ng|M|n'e)(n'e|M|ng) \times \exp\{it[\mathscr{H}'(n'e) - \mathscr{H}'(n, g)]\} \\ &= B \sum_n P(n) \sum_{n'} (n|M_{ge} \exp[it\mathscr{H}(e)]|n') \times (n'|M_{eg} \exp[-it\mathscr{H}(g)]|n). \end{aligned} \qquad (6.4)$$

Here $\mathscr{H}(g)$ is the total Hamiltonian of the system with the electron in its ground state. It is obtained by adding equations (5.3) and (5.5); $\mathscr{H}(e)$ is the same operator, but with the electron in the excited state. We have introduced the notation

$$M_{eg} = (e|M|g). \qquad (6.5)$$

Remembering the completeness relation for the final vibrational states, we have

$$C_{ge}(t) = B \sum_n P(n)(n|M_{ge} \exp[it\mathscr{H}(e)] M_{eg} \exp[-it\mathscr{H}(g)]|n). \qquad (6.6)$$

Thus the first few moments are

$$\langle \omega^0 \rangle = B \sum_n P(n)(n|M_{eg}^+ M_{eg}|n). \qquad (6.7)$$

$$\langle \omega \rangle = B \sum_n P(n)\{(n|M_{eg}^+ M_{eg}(\mathscr{H}(e) - \mathscr{H}(g))|n) + (n|M_{eg}^+[\mathscr{H}(e), M_{eg}]|n)\}. \qquad (6.8)$$

$$\langle \omega^2 \rangle = B \sum_n P(n)\{(n|M_{eg}^+ M_{eg}(\mathscr{H}(e) - \mathscr{H}(g))^2|n) + (n|M_{eg}^+[\mathscr{H}^2(e), M_{eg}] - 2M_{eg}^+[\mathscr{H}(e), M_{eg}]\mathscr{H}(g) + M_{eg}^+ M_{eg}[\mathscr{H}(g), \mathscr{H}(e)]|n)\}. \qquad (6.9)$$

If the Condon approximation is made, that is, if the electric dipole matrix element M_{eg} is assumed to be independent of the nuclear coördinates, then the moments become

$$\langle \omega^0 \rangle = B|M_{eg}|^2 \qquad (6.10)$$

$$\langle \omega \rangle = B|M_{eg}|^2 \sum_n P(n)(n|\mathscr{H}(e) - \mathscr{H}(g)|n) \qquad (6.11)$$

$$\langle \omega^2 \rangle = B|M_{eg}|^2 \sum_n P(n)\{(n|(\mathscr{H}(e) - \mathscr{H}(g))^2|n) + (n|[\mathscr{H}(g), \mathscr{H}(e)]|n)\}. \qquad (6.12)$$

VOL. 13 1960

The spectrum can then be normalized by choosing

$$B = |M_{eg}|^{-2}. \tag{6.13}$$

The zeroth and the first moments agree with the moments of the semiclassical spectrum as discussed by LAX. The first term on the right-hand side of equation (6.12) also agrees with the semiclassical result for the second moment, but the second term is a small quantum-mechanical correction, which will later be shown to vanish in first order. We shall now proceed with the evaluation of equation (6.11) and (6.12). Later we shall investigate the effects of the Condon approximation more fully.

Let us first write a formal expression for the transition energy operator. Thus,

$$\mathscr{H}(e)-\mathscr{H}(g) = \Delta_0+\sum_\lambda Q_\lambda\Delta_\lambda+\tfrac{1}{2}\sum_\mu\sum_\lambda \Delta_{\mu\lambda}Q_\mu Q_\lambda. \tag{6.14}$$

Denoting the thermal vibrational average over the initial states by Av( ) and the quantum-mechanical average over the initial vibrational wave functions by a bar, we have, with $(Q^0)_{\text{ground state}} = Q_\lambda^{(g)}$,

$$\langle\omega\rangle = \Delta_0+\sum_\lambda\Delta_\lambda Q_\lambda^{(g)}+\tfrac{1}{2}\sum_\mu\sum_\lambda\Delta_{\mu\lambda}\mathrm{Av}(\overline{Q_\mu Q_\lambda}) \tag{6.15}$$

$$\langle\omega^2\rangle = \Delta_0^2+\sum_\mu\sum_\lambda\Delta_\mu\Delta_\lambda\mathrm{Av}(\overline{Q_\mu Q_\lambda})+2\Delta_0\sum_\lambda\Delta_\lambda Q_\lambda^{(g)}+ +\Delta_0\sum_\mu\sum_\lambda\Delta_{\mu\lambda}\mathrm{Av}(\overline{Q_\mu Q_\lambda})+\mathrm{Av}\{\overline{[\mathscr{H}(g),\mathscr{H}(e)]}\} \tag{6.16}$$

and

$$\begin{aligned}\langle\omega^2\rangle-(\langle\omega\rangle)^2 &= \sum_\mu\sum_\lambda\Delta_\mu\Delta_\lambda\{\mathrm{Av}(\overline{Q_\mu Q_\lambda})-Q_\mu^{(g)}Q_\lambda^{(g)}\}+\mathrm{Av}\{\overline{[\mathscr{H}(g),\mathscr{H}(e)]}\}\\ &= \sum_\mu\sum_\lambda\Delta_\mu\Delta_\lambda\mathrm{Av}\{\overline{(Q_\lambda-Q_\lambda^{(g)})(Q_\mu-Q_\mu^{(g)})}\}+\mathrm{Av}\{\overline{[\mathscr{H}(g),\mathscr{H}(e)]}\}\\ &= \sum_\mu\Delta_\mu^2\mathrm{Av}[\overline{(Q_\mu-Q_\mu^{(g)})^2}]+\mathrm{Av}\{\overline{[\mathscr{H}(g),\mathscr{H}(e)]}\}.\end{aligned} \tag{6.17}$$

Thus it is clear that we could break off equation (6.15) after the first two terms, and equation (6.17) would remain unchanged. This is indeed the accuracy to which we have performed our expansions in the earlier part of the calculation, since we had assumed that the vibrational frequencies of the system were unaffected by the electronic transition. To this accuracy, it can also be readily seen that the second term in equation (6.17) vanishes. We thus have the following simplified expressions for the mean frequency and the variance, where we have carried out the averaging operations:

$$\langle\omega\rangle = \Delta_0+\sum_\lambda\Delta_\lambda Q_\lambda^{(g)} \tag{6.18}$$

$$\langle\omega^2\rangle-(\langle\omega\rangle)^2 = \sum_\lambda\Delta_\lambda^2\omega_\lambda^{-1}(n_\lambda+\tfrac{1}{2}). \tag{6.19}$$

The omission of the third term of equation (6.15) has the effect of making $\langle\omega\rangle$ independent of the temperature.

It now only remains to carry out the summations over the normal modes. For this purpose it is necessary to put these sums into the familiar form (4.29) for which we have obtained a summation formula in Section 4. We note that

$$\begin{aligned}\Delta_\lambda &\equiv \frac{\partial}{\partial Q_\lambda}[\mathscr{H}(e)-\mathscr{H}(g)]|_{Q_\lambda=0}\\ &= \sum_n\frac{\partial\xi_n}{\partial Q_\lambda}\frac{\partial}{\partial\xi_n}[\mathscr{H}(e)-\mathscr{H}(g)]|_{\xi_n=0}\\ &= \sum_n\Delta_n\frac{\partial\xi_n}{\partial Q_\lambda} = M^{-1/2}\sum_n\Delta_n W_n^\lambda\end{aligned} \tag{6.20}$$

where we used (5.2), and where Δ_n is defined by this equation. Using equation (5.4) and denoting E_n for the ground state by $E_n^{(g)}$, we have

$$\begin{aligned}\langle\omega\rangle &= \Delta_0-(1/M)\sum_m{}'\sum_n{}' E_m^{(g)}\Delta_n\sum_\lambda w_n^\lambda w_m^\lambda\omega_\lambda^{-2}\\ &= \Delta_0+\sum_m{}'\sum_n{}' E_m^{(g)}\Delta_n\Gamma_{mn}(0)\end{aligned} \tag{6.21}$$

where we have utilized equation (4.29) and (4.34). For the special defect discussed in section 4,

$\Gamma_{mn}(0)$ is given by equations (4.36) and (4.37). The series (6.21) converges exponentially, and it can therefore be summed numerically.

The analogous series for the variance can be summed directly only in the high-temperature limit, where

$$2n_\lambda+1 = \coth(\omega_\lambda/2k_BT) \approx 2k_BT/\omega_\lambda. \quad (6.22)$$

Then,

$$\begin{aligned}\langle\omega^2\rangle-(\langle\omega\rangle)^2 &= k_BT\sum_\lambda \Delta_\lambda^2\omega_\lambda^{-2}\\ &= k_BTM^{-1}\sum_m{}'\sum_n{}'\Delta_m\Delta_n\sum_\lambda w_m^\lambda w_n^\lambda\omega_\lambda^{-2}\\ &= -k_BT\sum_m{}'\sum_n{}'\Delta_m\Delta_n\Gamma_{mn}(0).\end{aligned} \quad (6.23)$$

This series also converges exponentially. This can be seen from an examination of the explicit form of Δ_n. From its definition (6.20), and from the fact that the lattice alone (that is, aside from the influence of the trapped electron) is in equilibrium, it is clear that all linear terms in the expansion of $\mathscr{H}(e)-\mathscr{H}(g)$ must originate from the electronic energy alone (see equation (5.1)), and therefore

$$\Delta_n = E_n^{(e)}-E_n^{(g)}. \quad (6.24)$$

Equations (3.16a) and (3.16b) show that these quantities do indeed fall off exponentially with n.

We can now also examine the effects of a departure from the Condon approximation. If the electric dipole matrix element has the form

$$M_{eg} = M^{(0)}+\sum_\lambda M^{(\lambda)}Q_\lambda, \quad (6.25)$$

then to first order in the displacements the mean absorption frequency can be shown to remain unchanged. The details of the effect on the variance have not been computed, but it appears that there will be a change.

We shall now turn to a discussion of the emission spectrum. It is given by a formula analogous to equation (6.3), but with the statistical weighting factor applied to the n' summation, that is, to the initial vibrational states, which in this case are the vibrational states when the electron is in the first excited state. The analogue of the equation (6.6) for the characteristic function is now

$$C_{eg}(t) = B'\sum_n P(n')(n'|M_{eg}\exp[-it\mathscr{H}(g)]M_{ge}\exp[it\mathscr{H}(e)]|n'). \quad (6.26)$$

It differs from equation (6.6) only in having the summation run over n' and in having $\mathscr{H}(g)$ replaced by $-\mathscr{H}(e)$, and $\mathscr{H}(e)$ replaced by $-\mathscr{H}(g)$. Formulas (6.7) through (6.14) can thus be taken over from the previous discussion, providing that the above replacements are made. Formulas (6.15) through (6.19) can also be taken over, provided that we remember that the quantum-mechanical averaging, which is indicated by a bar, is now over the excited electronic state vibrational states, and consequently the $Q_\lambda^{(g)}$'s must be replaced by $Q_\lambda^{(e)}$. Similarly, in equation (6.21) $E_m^{(g)}$ must be replaced by $E_m^{(e)}$. Thus we can obtain an expression for the shift of the mean frequency between the emission and the absorption. It is

$$\sum_m\sum_n\Delta_m\Delta_n\Gamma_{mn}(0). \quad (6.27)$$

7. DISCUSSION

The above calculations indicate that both the localized and the extended lattice-vibration modes play an important part in determining the shape of the absorption and the emission spectra. This is clearly seen from an examination of equation (6.18) as well as from a scrutiny of equation (6.19). Both the mean frequency and the variance are given by sums. From equations (5.4) and (4.15) it is clear that $Q_\lambda^{(g)}$ is of order unity for a localized mode, and of order $N^{-1/2}$ for an extended mode. Equation (6.20) shows that Δ_λ is also of order unity for a localized mode, and of order $N^{-1/2}$ for an extended mode. Since there are only several localized modes, but approximately N extended modes, it should be clear that all the extended modes combine to give a non-zero contribution to the sum in equation (6.18). Each localized mode also gives a non-vanishing contribution. An analogous statement clearly also holds for equation (6.19). This conclusion cannot be modified by the retention of the second-order terms in the expansions of the electronic energy in equations (3.16a) and (3.16b), since an equation of the form of (6.14) will still hold. Thus it will be necessary in general to treat deep-trap problems by a technique which can handle both the extended vibration modes and the localized modes. In special cases it is quite possible, of course, that the contribution of one or the

other kind of vibration modes should be dominant, but this cannot be taken for granted in general.

Approximate methods designed to handle both kinds of vibration modes may be tested by applying them to the present model. This comparison should bring out the influence of the assumption of an incorrect vibration spectrum (as in the Debye approximation) on the absorption and emission spectra of an electron. Under some special circumstances, this calculation may also provide a clue to the accuracy of an approximate method used for computing the localized modes.

It appears to the authors that the method used in the present paper for the calculation of the electronic wave function may be capable of extension to three dimensions and more realistic potentials. The MONTROLL and POTTS technique can probably also be used to handle the three dimensional problem. It would be interesting, therefore, to attempt to treat a three-dimensional problem by an extension of the present method.

VOL. 13 1960

Acknowledgements—The authors would like to thank Dr. HUBERT M. JAMES for his stimulating comments. They are also indebted to Mrs. ARDITH I. BAILEY for typing and editorial assistance. One of us (B.S.G.) would like to thank Dr. F. T. MCCLURE for his continued support and encouragement.

APPENDIX A

In this Appendix we carry out the inversion of the equations (3.9) to obtain α explicitly as a function of ξ_n and ξ_{-n}. The energy of the electron, E, is related to α by

$$E = -\tfrac{1}{2}\alpha^2. \tag{A.1}$$

The equation (3.10) can be written as

$$f(\alpha) = f_0(\alpha)+f_1(\alpha) = 0 \tag{A.2}$$

where $f_0(\alpha)$ does not depend explicitly on ξ_n and ξ_{-n}, while $f_1(\alpha)$ has an explicit linear dependence on ξ_n and ξ_{-n}, and so on. We wish to write

$$E = -\frac{\alpha^2}{2} = E_0+ + \sum_{n=1}^{\infty}\left\{\left(\frac{\partial E}{\partial \xi_n}\right)_0\xi_n+\left(\frac{\partial E}{\partial \xi_{-n}}\right)_0\xi_{-n}\right\}+\cdots \tag{A.3}$$

where the subscript 0 indicates the situation with all ξ_n and ξ_{-n} set equal to zero. However, we see that

$$\left(\frac{\partial E}{\partial \xi_n}\right)_0 = -\frac{1}{2}\frac{\partial}{\partial \xi_n}(\alpha^2)\Big|_0 = -\alpha\frac{\partial \alpha}{\partial \xi_n}\Big|_0, \tag{A.4}$$

so that we need $\partial\alpha/\partial\xi_n|_0$ and $\partial\alpha/\partial\xi_{-n}|_0$. These can be obtained by implicit differentiation of equation (A.2):

$$0 = f_0'(\alpha)\frac{\partial \alpha}{\partial \xi_n}+f_1'(\alpha)\frac{\partial \alpha}{\partial \xi_n}+\frac{\partial f_1(\alpha)}{\partial \xi_n} \tag{A.5}$$

with a similar expression from the differentiation with respect to ξ_{-n}. The primes denote differentiation with respect to the argument α.

Thus we find that

$$\frac{\partial \alpha}{\partial \xi_n} = -\frac{\partial f_1(\alpha)/\partial \xi_n}{f_0'(\alpha)+f_1'(\alpha)} \tag{A.6}$$

and

$$\left(\frac{\partial \alpha}{\partial \xi_n}\right)_0 = -\frac{1}{f_0'(\alpha_0)}\left(\frac{\partial f_1(\alpha_0)}{\partial \xi_n}\right)_0. \tag{A.7}$$

The quantity $E_0 = -\alpha_0^2/2$ is obtained from the solution to the equation

$$f_0(\alpha_0) = 0, \tag{A.8}$$

which in general must be carried out numerically. α_0 is the root of (A.8).

The explicit expressions for the functions $f_0(\alpha)$ and $f_1(\alpha)$ in the cases of the electron in its ground and excited states are:

(a) *Ground State*:

$$f_0(\alpha) = -\frac{1}{U_0}+\frac{1}{\alpha}[1+\exp(-2\alpha a)]- -\frac{2}{\alpha}\frac{\cosh 2\alpha a+1}{\left(1+\frac{\alpha}{v}\right)[\exp(2\alpha b)-1]}$$

$$f_1(\alpha) = \frac{2(\cosh 2\alpha a+1)}{1+\frac{\alpha}{v}}\sum_{n=1}^{\infty}\exp(-2n\alpha b)(\xi_n-\xi_{-n}) \tag{A.9}$$

(b) *Excited State*:

$$f_0(\alpha) = -\frac{1}{U_0}+\frac{1}{\alpha}[1-\exp(-2\alpha a)]- -\frac{2}{\alpha}\frac{\cosh 2\alpha a-1}{\left(1+\frac{\alpha}{v}\right)[\exp(2\alpha b)-1]}$$

$$f_1(\alpha) = \frac{2(\cosh 2\alpha a - 1)}{1+\dfrac{\alpha}{v}} \sum_{n=1}^{\infty} \exp(-2n\alpha b)(\xi_n - \xi_{-n}) \tag{A.10}$$

APPENDIX B

The explicit form of the Green's function for the perfect lattice is obtained as follows. We note that the Kronecker delta can be written as

$$\delta_{mk} = (1/N) \sum_{s=0}^{N-1} \exp[2\pi i s(m-k)/N]. \tag{B.1}$$

We therefore assume a form

$$\gamma_{mk}(\omega) = (1/N) \sum_{s=0}^{N-1} D_s \exp[2\pi i s(m-k)/N], \tag{B.2}$$

and substitute it into equation (4.9) in order to determine D_s. The equation (4.9) is satisfied if

$$D_s = -[-M\omega^2 + k_1 + 2k_2\{1 - \cos(2\pi s/N)\}]^{-1}. \tag{B.3}$$

This can also be readily rewritten as

$$\gamma_{mk}(\omega) = (1/N) \times \times \sum_{s=0}^{N-1} \frac{-\cos\{2\pi s(m-k)/N\}}{-M\omega^2 + k_1 + 2k_2\{1 - \cos(2\pi s/N)\}}. \tag{B.4}$$

Clearly this Green's function is a function of $|m-k|$ only.

If $M\omega^2 < k_1$, or if $M\omega^2 > k_1 + 4k_2$, then it is permissible to replace (B.4) by an integral obtained by setting $\theta = 2\pi s/N$ and proceeding to the limit as N becomes very large. This leads to the representation

$$\gamma_{mk}(\omega) = (-\tfrac{1}{2}k_2)(1/2\pi) \times \times \int_0^{2\pi} (p - \cos\theta)^{-1} \cos\{(m-k)\theta\}\, d\theta, \tag{B.5}$$

where p is given by

$$p = 1 + (k_1/2k_2) - (M\omega^2/2k_2). \tag{B.6}$$

This integral can be carried out by contour integration. Two cases must be distinguished. Case I arises when the localized mode is above the band. Then $M\omega^2 > k_1 + 4k_2$, and $p < -1$. If the localized mode is below the allowed band, however, we have case II, that is, $M\omega^2 < k_1$, and $p > 1$. For case I, the contour integration gives with $n = |m-k|$

$$\gamma_n(\omega) = (2k_2)^{-1}(-1)^n \operatorname{csch}\theta \exp(-n\theta) \tag{B.7}$$

where we have set $p = -\cosh\theta$. For case II we obtain

$$\gamma_n(\omega) = -(2k_2)^{-1} \operatorname{csch}\theta \exp(-n\theta), \tag{B.8}$$

where we set $p = \cosh\theta$.

REFERENCES

1. PEKAR S. I., *Untersuchungen über die Elektronentheorie der Kristalle*. Akademie-Verlag, Berlin (1954).
2. HUANG K. and RHYS A., *Proc. Roy. Soc.* A**204,** 406 (1950).
3. LAX M., *J. Chem. Phys.* **20,** 1752 (1952).
4. O'ROURKE R. C., *Phys. Rev.* **91,** 265 (1953).
5. MEYER H. J. G., *Physica* **21,** 253 (1955).
6. GOURARY B. S. and ADRIAN F. J., *Solid State Physics* (Ed. F. SEITZ and D. TURNBULL) Vol. 10. Academic Press, New York (1960).
7. SEITZ F., *Modern Theory of Solids* p. 281. McGraw-Hill, New York (1940).
8. MONTROLL E. W. and POTTS R. B., *Phys. Rev.* **100,** 525 (1955).

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 13. pp. 105–123. Printed in Great Britain.

THE DIAMAGNETISM OF CONDUCTION ELECTRONS IN METALS

J. E. HEBBORN and E. H. SONDHEIMER

Department of Mathematics, Queen Mary College, London

(*Received* 17 *August* 1959)

VOL. 13 1960

Abstract—The partition function is calculated to order H^2 for a gas of independent electrons interacting with a periodic lattice potential and a uniform applied magnetic field H. A formula is deduced for the steady diamagnetic susceptibility of a metal, expressed in terms of the Bloch wave functions and energy levels. The result is exact and general within the limits of the Bloch model.

1. INTRODUCTION

THE theory of the magnetic susceptibility of solids that arises from the translational motion of conduction electrons belongs to an important class of problems in solid-state physics that involve the energy levels and wave functions of the conduction electrons in the presence of an external magnetic field. For a free electron in a uniform field the solution of the Schrödinger equation is readily obtained,[1] but in real lattices the influence of the periodic potential must also be considered; the eigenvalue problem is then of a much greater order of difficulty and, in spite of important recent advances,[2] is even now incompletely solved. However, to determine the field-independent part of the magnetic susceptibility χ, with which we shall be concerned in the present paper, it is sufficient to calculate the partition function Z per unit volume (or, more generally, the density matrix) for the problem to order H^2 in the applied magnetic field H, and WILSON[3] has shown how to obtain the density matrix directly as a power series in H in terms of the solutions of the Schrödinger equation when $H = 0$. The calculation of χ is therefore essentially a computational problem, but in practice the computation is made so intractable by the appearance in large numbers of complicated interband matrix elements that no really satisfactory general expression for χ has been derived so far. In all existing treatments of the problem (including the present one), it is agreed that the result includes the well-known term χ_P, depending upon the Gaussian curvature of the energy surfaces, which was first derived by PEIERLS[4] for tightly bound Bloch electrons as a generalization of the free-electron susceptibility calculated by LANDAU[1], and it is also agreed that there must be a term in χ that reduces to the susceptibility of a set of free atoms in the limit when the lattice spacing tends to infinity. There are, however, additional terms in χ (with no simple analogues) which in previous treatments have been either incompletely evaluated[3] or appear as very complicated expressions;[5] it is only for the special case of electrons confined to the vicinity of the bottom or top of an energy band that it has so far been possible to derive a theoretical expression for χ in a form sufficiently simple to permit of numerical evaluation (KJELDAAS and KOHN[6]).

In the present paper we re-calculate χ for Bloch electrons by a method that is essentially identical with WILSON's[3]. The calculation is restricted to non-degenerate energy bands, and we ignore all effects associated with the electron spin. These restrictions could be lifted, but are unlikely to affect the results significantly except in special cases. We also neglect all electron–electron interaction terms, in the hope (for which there is some theoretical support) that their effects are small so long as we disregard the phenomenon of superconductivity (compare Ref. 7). Subject to these assumptions, we do not neglect any of the terms in the complete partition function (it is somewhat simpler to calculate Z directly instead of the full

density matrix), and we obtain the final general expression for χ (equations (53) to (56)), expressed in terms of the Bloch wave functions and energy levels, in a much simpler form than any that have been given before and one which, we hope, is not too far from being the simplest that can be obtained for the general case of arbitrary crystal symmetry and arbitrary orientations of H.

The formula for χ is obtained after reduction of a very complicated expression for Z and assumes its final, reasonably simple, form through the aid of apparently fortuitous cancellations of terms. This suggests that it should be possible to discover simpler formulations of the calculation, a problem of particular interest with a view to attempts at extending the theory to deal with phenomena such as the de Haas–van Alphen effect (see Ref. 8 and Note added in proof, p. 123).

The details of the calculation of Z are described in Sections 2–4, and the susceptibility formula is written down in Section 5. In Section 5(a) we discuss the invariance of the result under changes in the phase of the Bloch functions and in the gauge of the vector potential, and in Section 5(b) we show that the result reduces to the expected forms in the limiting cases of free and tightly bound electrons and in the case considered by KJELDAAS and KOHN[6]. In Section 5(c), finally, we add some remarks on the evaluation of the susceptibility, but no attempt is made in the present paper to derive numerical estimates of χ for particular substances.

2. THE PARTITION FUNCTION

We consider a system of electrons interacting with a periodic lattice potential $\mathscr{V}(\boldsymbol{r})$ and a constant external magnetic field H. We take H to be in the z-direction, choose the vector potential in the form $(-Hy, 0, 0)$ and write the one-electron Hamiltonian $\mathscr{H}$ as

$$\mathscr{H} = \mathscr{H}_0+\mathscr{H}_1+\mathscr{H}_2, \tag{1}$$

where

$$\mathscr{H}_0 = -\tfrac{1}{2}\nabla^2+\mathscr{V}, \qquad \mathscr{H}_1 = 2i\mu_0 Hy\frac{\partial}{\partial x},$$
$$\mathscr{H}_2 = 2(\mu_0 H)^2y^2 \tag{2}$$

and $\mu_0 = 1/2c$ is the Bohr magneton. (Atomic units are used, with $e = \hbar = m = 1$.)

The partition function of the system (volume V) is (WILSON[10], Section 6.61)

$$VZ = \mathrm{Tr}\{\exp(-\beta\mathscr{H})\} = \int \Psi(\boldsymbol{r}, \boldsymbol{r}, \beta)\, d\boldsymbol{r}, \tag{3}$$

where $\beta = 1/kT$ and

$$\Psi(\boldsymbol{r}, \boldsymbol{r}, \beta) = \sum_i \psi_i(\boldsymbol{r})^* \exp(-\beta\mathscr{H})\psi_i(\boldsymbol{r}), \tag{4}$$

defined with reference to any complete orthogonal set of functions ψ_i. Direct calculation of the expression (3) for an infinitely large system leads to difficulties, since, when the integration in equation (3) is extended over the whole of space, the result (which is proportional to the volume of the system) is infinite. In fact we require Z, the partition function per unit volume, and it can be shown from the transformation properties of $\Psi(\boldsymbol{r}, \boldsymbol{r}, \beta)$ under a change in the origin of co-ordinates[8] that the integral of $\Psi(\boldsymbol{r}, \boldsymbol{r}, \beta)$ over one unit cell of the lattice is the same for all lattice cells. Hence

$$Z = \frac{1}{\Delta}\int \Psi(\boldsymbol{r}, \boldsymbol{r}, \beta)\, d\tau_0, \tag{5}$$

where Δ is the volume of the unit cell and the integration extends over one unit cell. The quantity (5) is finite and can be calculated unambiguously without having to consider explicitly the influence of the boundaries of the system.

To calculate Z to order H^2 in the magnetic field we expand $\exp(-\beta\mathscr{H})$ in powers of $\mathscr{H}_1$ and $\mathscr{H}_2$ by means of the well-known operator expansion[11]

$$\exp[-(a+b)] = e^{-a} - \int_0^1 ds\, \exp[-(1-s)a]be^{-sa} +$$
$$+ \int_0^1 s\, ds \int_0^1 ds'\, \exp[-(1-s)a]b \times$$
$$\times \exp[-s(1-s')a]b \exp(-ss'a) - \ldots \tag{6}$$

The expansion of Z in powers of H is then

$$Z = Z_0+Z_2+Z_{1,1}+\ldots, \tag{7}$$

where Z_0 is the partition function when $H = 0$, and where

$$Z_2 = -\frac{\beta}{\Delta}\sum_i \int d\tau_0 \int_0^1 ds\psi_i^* \times$$
$$\times \exp[-(1-s)\beta\mathscr{H}_0]\mathscr{H}_2 \exp(-s\beta\mathscr{H}_0)\psi_i \tag{8}$$

and

$$Z_{1,1} = \frac{\beta^2}{\Delta} \sum_i \int d\tau_0 \int_0^1 s\, ds \int_0^1 ds' \psi_i^* \times$$

$$\times \exp[-(1-s)\beta\mathscr{H}_0]\mathscr{H}_1 \exp[-s(1-s')\beta\mathscr{H}_0]\mathscr{H}_1 \times$$
$$\times \exp(-ss'\beta\mathscr{H}_0)\psi_i. \quad (9)$$

These are the terms in Z proportional to H^2 that give rise to the steady diamagnetic susceptibility.†

To evaluate the integrals (8) and (9) we use as basis functions ψ_i the Bloch functions ψ_{km}, expressed in terms of the reduced wave vector $\boldsymbol{k}$ (with components k_1, k_2, k_3) and a band index m. These functions are eigenfunctions of the zero-field Hamiltonian $\mathscr{H}_0$ with eigenvalues E_{km}, and they form a complete set of orthogonal functions for all space. We suppose them normalized so that

$$\frac{\Delta}{(2\pi)^3} \int \psi_{km}^*(\boldsymbol{r})\psi_{kn}(\boldsymbol{r})\, d\boldsymbol{r} = \delta(\boldsymbol{k}-\boldsymbol{k}')\delta_{mn} \quad (10)$$

(orthogonality), and

$$\frac{\Delta}{(2\pi)^3} \sum_m \int \psi_{km}^*(\boldsymbol{r}')\psi_{km}(\boldsymbol{r})\, d\boldsymbol{k} = \delta(\boldsymbol{r}-\boldsymbol{r}') \quad (11)$$

(completeness). The integration with respect to $\boldsymbol{k}$ extends over the fundamental Brillouin zone. If we write, as usual,

$$\psi_{km}(\boldsymbol{r}) = u_{km}(\boldsymbol{r}) \exp(i\boldsymbol{k}\cdot\boldsymbol{r}), \quad (12)$$

the functions u_{km} possess the lattice periodicity and, for fixed $\boldsymbol{k}$, they form a complete orthonormal set in the unit cell, so that

$$\int u_{km}^*(\boldsymbol{r})u_{kn}(\boldsymbol{r})\, d\tau_0 = \delta_{mn},$$

and

$$\sum_m u_{km}^*(\boldsymbol{r}')u_{km}(\boldsymbol{r}) = \delta(\boldsymbol{r}-\boldsymbol{r}') \quad (13)$$

for $\boldsymbol{r}'$, $\boldsymbol{r}$ within the same cell. A number of relations involving Bloch functions, required in the subsequent analysis, are collected together in Appendix 1.

Equation (8) may now be rewritten as

$$Z_2 = -\frac{\beta}{(2\pi)^3} \sum_m \int d\boldsymbol{k} \int d\tau_0 \int_0^1 ds \psi_{km}^* \times$$

$$\times \exp[-(1-s)\beta\mathscr{H}_0]\mathscr{H}_2 \exp(-s\beta\mathscr{H}_0)\psi_{km},$$

or, in view of (11),

$$Z_2 = -\frac{\beta\Delta}{(2\pi)^6} \sum_{mn} \int d\boldsymbol{k} \int d\boldsymbol{k}' \int d\tau_0 \int d\boldsymbol{r}' \int_0^1 ds \psi_{km}^*(\boldsymbol{r}) \exp[-(1-s)\beta\mathscr{H}_0] \times$$

$$\times \psi_{kn}(\boldsymbol{r})\psi_{kn}^*(\boldsymbol{r}')\mathscr{H}_2(\boldsymbol{r}') \exp(-s\beta\mathscr{H}_0)\psi_{km}(\boldsymbol{r}') = -\frac{2\beta\Delta}{(2\pi)^6}(\mu_0 H)^2 \sum_{mn} \int d\boldsymbol{k} \int d\boldsymbol{k}' \int d\tau_0 \int d\boldsymbol{r}' \int_0^1 ds \times$$

$$\times \exp\{-(1-s)\beta E_{kn} - s\beta E_{km}\}\psi_{km}^*(\boldsymbol{r})\psi_{kn}(\boldsymbol{r})\psi_{kn}^*(\boldsymbol{r}')y'^2\psi_{km}(\boldsymbol{r}'),$$

since $\mathscr{H}_0\psi_{km} = E_{km}\psi_{km}$ and $\mathscr{H}_2(\boldsymbol{r}) = 2(\mu_0 H)^2 y^2$.

Since Z is real, we may replace this expression by its complex conjugate. It is then in a form in which we can apply formula (A.2) of Appendix 1 to perform the integrations with respect to $\boldsymbol{r}'$ and $\boldsymbol{k}'$. This gives

$$Z_2 = -\frac{2\beta}{(2\pi)^3}(\mu_0 H)^2 \sum_{mn} \int d\boldsymbol{k} \int d\tau_0 \int_0^1 ds \psi_{km} \exp(-s\beta E_{km}) L_{kmn}^{(2)} \exp\{-(1-s)\beta E_{kn}\}\psi_{kn}^*, \quad (14)$$

† The expansion (6) also gives rise to a term in Z which is linear in H. This term is easily shown to reduce to zero, and has been omitted.

where

$$L^{(2)}_{kmn} = -\delta_{mn}\frac{\partial^2}{\partial k_2^2} - 2\left(\int u^*_{km}\frac{\partial u_{kn}}{\partial k_2}\,d\tau_0\right)\frac{\partial}{\partial k_2} - \int u^*_{km}\frac{\partial^2 u_{kn}}{\partial k_2^2}\,d\tau_0. \tag{15}$$

An analogous calculation, which involves a double application of formula (A.1) of Appendix 1, gives

$$Z_{1,1} = \frac{4\beta^2}{(2\pi)^3}(\mu_0 H)^2 \sum_{lmn}\int d\boldsymbol{k}\int d\tau_0 \int_0^1 s\,ds \int_0^1 ds'\psi_{kl}\exp(-ss'\beta E_{kl})\times$$

$$\times L^{(1,1)}_{klm}\exp\{-s(1-s')\beta E_{km}\}L^{(1,1)}_{kmn}\exp\{-(1-s)\beta E_{kn}\}\psi^*_{kn}, \tag{16}$$

where

$$L^{(1,1)}_{kmn} = \left(\int \psi^*_{km}\frac{\partial\psi_{kn}}{\partial x}\,d\tau_0\right)\frac{\partial}{\partial k_2} + ik_1\int u^*_{km}\frac{\partial u_{kn}}{\partial k_2}\,d\tau_0 - \int \frac{\partial u^*_{km}}{\partial x}\frac{\partial u_{kn}}{\partial k_2}\,d\tau_0. \tag{17}$$

In equations (14) and (16) the differential operators L_{kmn} operate on all terms standing to the right.

The partition function has now been expressed in the form of an integral over the fundamental Brillouin zone in $\boldsymbol{k}$-space, in which the integrand involves the Bloch energy levels E_{km} and wave functions ψ_{km} for the different bands m at a fixed value of $\boldsymbol{k}$. In future we omit explicit reference to the wave vector $\boldsymbol{k}$.

3. REDUCTION OF Z_2

From equations (14) and (15) we obtain, performing the differentiations with respect to k_2,

$$Z_2 = \frac{2\beta}{(2\pi)^3}(\mu_0 H)^2 \sum_{mn}\int d\boldsymbol{k}\int d\tau_0\psi_m\left[\left\{-\delta_{mn}\beta f_2(\epsilon_{mn})\frac{\partial^2 E_n}{\partial k_2^2} + \delta_{mn}\beta^2 f_3(\epsilon_{mn})\left(\frac{\partial E_n}{\partial k_2}\right)^2 - \right.\right.$$

$$-2\beta f_2(\epsilon_{mn})\frac{\partial E_n}{\partial k_2}\int u^*_m\frac{\partial u_n}{\partial k_2}\,d\tau_0 + f_1(\epsilon_{mn})\int u^*_m\frac{\partial^2 u_n}{\partial k_2^2}\,d\tau_0\Big\}\psi_n^* + \Big\{-2\delta_{mn}\beta f_2(\epsilon_{mn})\frac{\partial E_n}{\partial k_2} +$$

$$\left.\left. + 2f_1(\epsilon_{mn})\int u^*_m\frac{\partial u_n}{\partial k_2}\,d\tau_0\Big\}\frac{\partial\psi_n^*}{\partial k_2} + \delta_{mn}f_1(\epsilon_{mn})\frac{\partial^2\psi_n^*}{\partial k_2^2}\right]\exp(-\beta E_n), \tag{18}$$

where $\epsilon_{mn} = \beta(E_n - E_m)$, and

$$f_r(x) = \int_0^1 (1-s)^{r-1}e^{xs}\,ds$$

$$= \frac{(r-1)!}{x^r}\left\{e^x - \sum_{P=0}^{r-1}\frac{x^P}{P!}\right\} \quad (r = 1, 2, 3). \tag{19}$$

Note that $f_r(\epsilon_{mm}) = f_r(0) = 1/r$.

Since

$$\int \psi_m\psi_n^*\,d\tau_0 = \delta_{mn},$$

all the terms in Z_2, except the second term involving $\partial\psi_n^*/\partial k_2$, contain a factor δ_{mn}, so that for these the summation over n can be carried out at once. Among the terms thus obtained, consider in particular

$$\frac{2\beta}{(2\pi)^3}(\mu_0 H)^2 \sum_m \int d\boldsymbol{k}\left\{-\beta\frac{\partial E_m}{\partial k_2}\int u^*_m\frac{\partial u_m}{\partial k_2}\,d\tau_0 + \right.$$

$$\left. + \int u^*_m\frac{\partial^2 u_m}{\partial k_2^2}\,d\tau_0\right\}\exp(-\beta E_m). \tag{20}$$

Since $u_{-km} = u^*_{km}$, the first term is the integral with respect to $\boldsymbol{k}$ of an odd function of $\boldsymbol{k}$, and is therefore zero. Further, since Z is real and

$$\frac{\partial^2}{\partial k_2^2}\int u^*_m u_m\,d\tau_0 = 0,$$

the second term may be written as

$$-\frac{2\beta}{(2\pi)^3}(\mu_0 H)^2 \sum_m \int d\boldsymbol{k} \int \left|\frac{\partial u_m}{\partial k_2}\right|^2 d\tau_0 \exp(-\beta E_m).$$

Two further terms in Z_2 are of the same form as (20), except that u_m is replaced by ψ_m^*, and are treated in the same way. Finally, it is convenient to rearrange the surviving double sum in Z_2 by collecting together all terms in the sum which involve a particular exponential factor $\exp(-\beta E_m)$ (say). We use the formula

$$\sum_{mn} f_1(\epsilon_{mn}) G_{mn} \exp(-\beta E_n) = \sum_m G_{mm} \exp(-\beta E_m) + \sum_{mn}^{m\neq n} G_{mn} \frac{\exp(-\beta E_m) - \exp(-\beta E_n)}{\beta(E_n - E_m)}$$

$$= \sum_m \left\{ G_{mm} + \sum_{n\neq m} \frac{G_{mn} + G_{nm}}{\beta(E_n - E_m)} \right\} \exp(-\beta E_m). \tag{21}$$†

When all the simplifications have been made, we are left with

$$Z_2 = \frac{2}{(2\pi)^3}(\mu_0 H)^2 \sum_m \int d\boldsymbol{k} \Big[-\tfrac{1}{2}\beta^2 \frac{\partial^2 E_m}{\partial k_2^2} + \tfrac{1}{3}\beta^3 \left(\frac{\partial E_m}{\partial k_2}\right)^2 -$$

$$-\beta \int \left|\frac{\partial u_m}{\partial k_2}\right|^2 d\tau_0 - \beta \int \left|\frac{\partial \psi_m}{\partial k_2}\right|^2 d\tau_0 + 2\beta \int \psi_m \frac{\partial \psi_m^*}{\partial k_2} d\tau_0 \int u_m^* \frac{\partial u_m}{\partial k_2} d\tau_0 +$$

$$+2 \sum_{n\neq m} \frac{1}{E_n - E_m} \Big\{ \int \psi_m \frac{\partial \psi_n^*}{\partial k_2} d\tau_0 \int u_m^* \frac{\partial u_n}{\partial k_2} d\tau_0 +$$

$$+ \int \psi_m^* \frac{\partial \psi_n}{\partial k_2} d\tau_0 \int u_m \frac{\partial u_n^*}{\partial k_2} d\tau_0 \Big\} \Big] \exp(-\beta E_m). \tag{22}$$

† $\sum_{mn}^{m\neq n}$ denotes the double summation $\sum_m \sum_n$ in which terms with $m = n$ are omitted. $\sum_{n\neq m}$ denotes a single summation over n, for fixed m.

4. REDUCTION OF $Z_{1,1}$

We introduce the notation

$$P_{mn} = \int \psi_m^* \frac{\partial \psi_n}{\partial x} d\tau_0,$$

$$Q_{mn} = ik_1 \int u_m^* \frac{\partial u_n}{\partial k_2} d\tau_0 - \int \frac{\partial u_m^*}{\partial x} \frac{\partial u_n}{\partial k_2} d\tau_0, \tag{23}$$

so that

$$L_{mn}^{(1,1)} = P_{mn} \frac{\partial}{\partial k_2} + Q_{mn}.$$

(P_{mn} is i times the matrix element between Bloch waves of the x-component of the momentum operator.) Then, performing the differentiations with respect to k_2 in (16), we find after some algebra that

$$Z_{1,1} = \frac{4\beta^2}{(2\pi)^3}(\mu_0 H)^2 \sum_{lmn} \sum_{i=1}^{3} \sum_{j=1}^{2} \int d\boldsymbol{k} I_{ij}(lmn) A_{ij}(lmn), \tag{24}$$

VOL. 13 1960

where

$$I_{ij}(lmn) = \exp(-\beta E_n) \int_0^1 s^i \exp[s\beta(E_n - E_m)]\, ds \int_0^1 s'^{j-1} \exp[ss'\beta(E_m - E_l)]\, ds', \tag{25}$$

and

$$A_{11}(lmn) = \int d\tau_0 \psi_l \left(Q_{lm} - \beta \frac{\partial E_n}{\partial k_2} P_{lm} + P_{lm} \frac{\partial}{\partial k_2}\right)\left(Q_{mn} - \beta \frac{\partial E_n}{\partial k_2} P_{mn} + P_{mn} \frac{\partial}{\partial k_2}\right)\psi_n^*, \tag{26}$$

$$A_{12}(lmn) = 0, \tag{27}$$

$$A_{21}(lmn) = \beta \int d\tau_0 \psi_l \left\{\left(Q_{lm} - \beta \frac{\partial E_n}{\partial k_2} P_{lm} + P_{lm} \frac{\partial}{\partial k_2}\right)\frac{\partial E_n}{\partial k_2} P_{mn} + \right.$$
$$\left. + \left(\frac{\partial E_n}{\partial k_2} - \frac{\partial E_m}{\partial k_2}\right) P_{lm}\left(Q_{mn} - \beta \frac{\partial E_n}{\partial k_2} P_{mn} + P_{mn} \frac{\partial}{\partial k_2}\right)\right\}\psi_n^*, \tag{28}$$

$$A_{22}(lmn) = \beta \int d\tau_0 \psi_l \frac{\partial E_m}{\partial k_2} P_{lm}\left(Q_{mn} - \beta \frac{\partial E_n}{\partial k_2} P_{mn} + P_{mn} \frac{\partial}{\partial k_2}\right)\psi_n^*, \tag{29}$$

$$A_{31}(lmn) = \beta^2 \int d\tau_0 \psi_l \left(\frac{\partial E_n}{\partial k_2} - \frac{\partial E_m}{\partial k_2}\right)\frac{\partial E_n}{\partial k_2} P_{lm} P_{mn} \psi_n^*, \tag{30}$$

$$A_{32}(lmn) = \beta^2 \int d\tau_0 \psi_l \frac{\partial E_m}{\partial k_2} \frac{\partial E_n}{\partial k_2} P_{lm} P_{mn} \psi_n^*. \tag{31}$$

The τ_0-integration in equations (30) and (31) gives a factor δ_{ln}; the other A-integrals contain, however, in addition to terms proportional to δ_{ln}, also terms proportional to

$$U_{ln} = \int \psi_l \frac{\partial \psi_n^*}{\partial k_2} d\tau_0 \text{ and } V_{ln} = \int \psi \frac{\partial^2 \psi_n^*}{\partial k_2^2} d\tau_0. \tag{32}$$

We now evaluate the elementary integrals (25) and then (as in Section 3) rearrange the terms in the triple summations over the band indices l, m, n in (24) so as to group together all terms involving a particular exponential factor $\exp(-\beta E_m)$.

For example, for $i = j = 1$ we obtain

$$I_{11}(lmn) = \frac{\exp(-\beta E_l)}{\beta^2(E_l - E_n)(E_l - E_m)} +$$
$$+ \frac{\exp(-\beta E_m)}{\beta^2(E_m - E_l)(E_m - E_n)} + \frac{\exp(-\beta E_n)}{\beta^2(E_n - E_m)(E_n - E_l)}, \tag{33}$$

with the special cases

$$I_{11}(mmn) = I_{11}(mnm) = I_{11}(nmm)$$
$$= \exp(-\beta E_m) f_2(\epsilon_{nm}) \tag{34}$$

(in the notation of Section 3), and

$$I_{11}(mmm) = \tfrac{1}{2} \exp(-\beta E_m). \tag{35}$$

We therefore have, using (19),

$$\sum_{lmn} I_{11}(lmn)A_{11}(lmn) = \sum_{lmn}^{l\neq m\neq n} I_{11}(lmn)A_{11}(lmn) + \sum_{mn}^{m\neq n} \{I_{11}(mmn)A_{11}(mmn)+$$

$$+I_{11}(mnm)A_{11}(mnm)+I_{11}(nmm)A_{11}(nmm)\}+\sum_{m} I_{11}(mmm)A_{11}(mmm)$$

$$=\sum_{m}\Bigg[\sum_{ln}^{l\neq n\neq m}\frac{1}{\beta^2(E_l-E_m)(E_n-E_m)}\{A_{11}(lmn)+A_{11}(mnl)+A_{11}(nlm)\}+$$

$$+\sum_{n\neq m}\frac{1}{\beta^2(E_n-E_m)^2}\{A_{11}(nnm)+A_{11}(nmn)+A_{11}(mnn)\}-$$

$$-\sum_{n\neq m}\frac{1-\beta(E_n-E_m)}{\beta^2(E_n-E_m)^2}\{A_{11}(mmn)+A_{11}(mnm)+A_{11}(nmm)\}+\tfrac{1}{2}A_{11}(mmm)\Bigg]\exp(-\beta E_m). \tag{36}$$

Note that the terms in $A_{11}(lmn)$ which are proportional to δ_{ln} give zero contribution to (36), except in the terms involving $A_{11}(nmn)$, $A_{11}(mnm)$ and $A_{11}(mmm)$.

When all the terms in (24) have been treated in this way, and the explicit expressions for the $A_{ij}(lmn)$ given by (26) to (31) are substituted, we obtain the expression for $Z_{1,1}$ in the form

$$Z_{1,1}=\sum_{m}\int d\boldsymbol{k}F_m(\boldsymbol{k})\exp(-\beta E_m), \tag{37}$$

where $F_m(\boldsymbol{k})$ is a sum of a large number of terms of the following form. $F_m(\boldsymbol{k})$ involves the derivatives of the energy with respect to k_1 and k_2, the integrals P_{mn} and Q_{mn} defined by (23) and their derivatives with respect to k_2, and the integrals U_{mn} and V_{mn} defined by (32). $F_m(\boldsymbol{k})$ contains single sums over the band index $n(n\neq m)$ with energy denominators $(E_n-E_m)^j$ for $j=1, 2, 3, 4$, and double sums over l and n $(l\neq n\neq m)$ with energy denominators $(E_l-E_m)(E_n-E_m)$ and $(E_l-E_m)^2(E_n-E_m)$. Finally, $F_m(\boldsymbol{k})$ is a polynomial in β of the fourth degree,

$$F_m(\boldsymbol{k})=\sum_{i=0}^{4}F_m^{(i)}(\boldsymbol{k})\beta^i. \tag{38}$$

It is convenient to group together terms in $Z_{1,1}$ that correspond to a particular power of β in $F_m(\boldsymbol{k})$. We therefore write

$$Z_{1,1}=\sum_{i=0}^{4}Z_{1,1}^{(i)},$$

where

$$Z_{1,1}^{(i)}=\beta^i\sum_{m}\int d\boldsymbol{k}F_m^{(i)}(\boldsymbol{k})\exp(-\beta E_m), \tag{39}$$

and, treating Z_2 in the same way,

$$Z_{1,1}+Z_2=\sum_{i=0}^{4}(Z_{1,1}^{(i)}+Z_2^{(i)})=\sum_{i=0}^{4}Z^{(i)}, \tag{40}$$

where $Z_2^{(i)}$ includes all terms proportional to $\beta^i\exp(-\beta E_m)$ in the integrand of (22).

The complete expression for $Z_{1,1}$ will not be written down here, and we shall discuss in detail the simplification of $Z^{(i)}$ for $i=4, 3, 2$ only. (In these terms there are no double sums over band indices.)

(a) *Reduction of* $Z^{(4)}+Z^{(3)}$

It is found that

$$Z^{(4)}+Z^{(3)}=\frac{2}{(2\pi)^3}(\mu_0 H)^2\beta^3\sum_m\int d\boldsymbol{k}\left[-\frac{1}{4}\beta\left(\frac{\partial E_m}{\partial k_1}\right)^2\left(\frac{\partial E_m}{\partial k_2}\right)^2+\frac{1}{3}\left(\frac{\partial E_m}{\partial k_2}\right)^2+\right.$$

$$+\frac{1}{3}\left(\frac{\partial E_m}{\partial k_1}\right)^2\frac{\partial^2 E_m}{\partial k_2^2}+\frac{1}{3}\frac{\partial E_m}{\partial k_1}\frac{\partial E_m}{\partial k_2}\frac{\partial^2 E_m}{\partial k_1\partial k_2}-i\frac{\partial E_m}{\partial k_1}\frac{\partial E_m}{\partial k_2}Q_{mm}+$$

$$\left.+\left(\frac{\partial E_m}{\partial k_1}\right)^2\frac{\partial E_m}{\partial k_2}U_{mm}+\frac{2}{3}\left(\frac{\partial E_m}{\partial k_2}\right)^2\sum_{n\neq m}\frac{1}{E_n-E_m}P_{mn}P_{nm}\right]\exp(-\beta E_m). \tag{41}$$

To simplify this we note that, since Z is real, any complex term in the expression for Z can be replaced by its real part. (The vanishing imaginary parts correspond to non-zero terms in the complete density matrix.) Therefore:

(i) the term in equation (41) involving U_{mm} is zero, since U_{mm} is purely imaginary (see equations (32) and (A.6));

(ii) in the term involving Q_{mm} we may replace Q_{mm} by

$$-\frac{1}{2}\int\left(\frac{\partial u_m^*}{\partial x}\frac{\partial u_m}{\partial k_2}-\frac{\partial u_m}{\partial x}\frac{\partial u_m^*}{\partial k_2}\right)d\tau_0$$

$$=\frac{1}{2}i\frac{\partial^2 E_m}{\partial k_1\partial k_2}$$

(see equations (23), (A.5) and (A.6));

(iii) using (A.3) and the completeness relation (13), we have

$$\sum_{n\neq m}\frac{1}{E_n-E_m}P_{mn}P_{nm}=i\sum_{n\neq m}\int u_n^*\frac{\partial u_m}{\partial x}d\tau_0\int u_m^*\frac{\partial u_n}{\partial k_1}d\tau_0$$

$$=-i\sum_{n\neq m}\int u_n^*\frac{\partial u_m}{\partial x}d\tau_0\int u_n\frac{\partial u_m^*}{\partial k_1}d\tau_0$$

$$=-i\left(\int\frac{\partial u_m}{\partial x}\frac{\partial u_m^*}{\partial k_1}d\tau_0-\int u_m^*\frac{\partial u_m}{\partial x}d\tau_0\int u_m\frac{\partial u_m^*}{\partial k_1}d\tau_0\right),$$

and the real part of this is, by (A.4),

$$-\frac{1}{2}i\int\left(\frac{\partial u_m}{\partial x}\frac{\partial u_m^*}{\partial k_1}-\frac{\partial u_m^*}{\partial x}\frac{\partial u_m}{\partial k_1}\right)d\tau_0=-\frac{1}{2}\left(1-\frac{\partial^2 E_m}{\partial k_1^2}\right) \tag{42}$$

(the imaginary part is zero).

Substituting these results into equation (41), we obtain

$$Z^{(4)}+Z^{(3)}=\frac{2}{(2\pi)^3}(\mu_0 H)^2\beta^3\sum_m\int d\boldsymbol{k}\left[-\frac{1}{4}\beta\left(\frac{\partial E_m}{\partial k_1}\right)^2\left(\frac{\partial E_m}{\partial k_2}\right)^2+\right.$$

$$\left.+\frac{1}{3}\left(\frac{\partial E_m}{\partial k_1}\right)^2\frac{\partial^2 E_m}{\partial k_2^2}+\frac{1}{3}\left(\frac{\partial E_m}{\partial k_2}\right)^2\frac{\partial^2 E_m}{\partial k_1^2}+\frac{5}{6}\frac{\partial E_m}{\partial k_1}\frac{\partial E_m}{\partial k_2}\frac{\partial^2 E_m}{\partial k_1\partial k_2}\right]\exp(-\beta E_m). \tag{43}$$

This expression can be transformed into a form which involves β only through a factor $\beta^2\exp(-\beta E_m)$, and which can then be combined later with the expression obtained for $Z^{(2)}$. We use the identity

$$\left\{-\frac{1}{4}\beta^2\left(\frac{\partial E}{\partial k_1}\right)^2\left(\frac{\partial E}{\partial k_2}\right)^2+\frac{1}{3}\beta\left(\frac{\partial E}{\partial k_1}\right)^2\frac{\partial^2 E}{\partial k_2^2}+\frac{1}{3}\beta\left(\frac{\partial E}{\partial k_2}\right)^2\frac{\partial^2 E}{\partial k_1^2}+\frac{5}{6}\beta\frac{\partial E}{\partial k_1}\frac{\partial E}{\partial k_2}\frac{\partial^2 E}{\partial k_1\partial k_2}\right\}\exp(-\beta E)$$

$$=\left\{\frac{5}{12}\frac{\partial^2 E}{\partial k_1^2}\frac{\partial^2 E}{\partial k_2^2}+\frac{1}{3}\left(\frac{\partial^2 E}{\partial k_1\partial k_2}\right)^2+\frac{3}{8}\left(\frac{\partial E}{\partial k_1}\frac{\partial^3 E}{\partial k_1\partial k_2^2}+\frac{\partial E}{\partial k_2}\frac{\partial^3 E}{\partial k_1^2\partial k_2}\right)\right\}\exp(-\beta E)+$$

$$+\frac{\partial}{\partial k_1}\left[\left\{\frac{1}{8}\beta\frac{\partial E}{\partial k_1}\left(\frac{\partial E}{\partial k_2}\right)^2-\frac{5}{24}\frac{\partial E}{\partial k_1}\frac{\partial^2 E}{\partial k_2^2}-\frac{1}{6}\frac{\partial E}{\partial k_2}\frac{\partial^2 E}{\partial k_1\partial k_2}\right\}\exp(-\beta E)\right]+$$

$$+\frac{\partial}{\partial k_2}\left[\left\{\frac{1}{8}\beta\frac{\partial E}{\partial k_2}\left(\frac{\partial E}{\partial k_1}\right)^2-\frac{5}{24}\frac{\partial E}{\partial k_2}\frac{\partial^2 E}{\partial k_1^2}-\frac{1}{6}\frac{\partial E}{\partial k_1}\frac{\partial^2 E}{\partial k_1\partial k_2}\right\}\exp(-\beta E)\right]. \quad (44)$$

On integrating this over the fundamental Brillouin zone, the complete derivatives with respect to k_1 and k_2 lead to surface integrals which vanish because of the periodicity of the integrands. We therefore have

$$Z^{(4)}+Z^{(3)}=\frac{2}{(2\pi)^3}(\mu_0 H)^2\beta^2\sum_m\int d\boldsymbol{k}\left[\frac{5}{12}\frac{\partial^2 E_m}{\partial k_1^2}\frac{\partial^2 E_m}{\partial k_2^2}+\right.$$

$$\left.+\frac{1}{3}\left(\frac{\partial^2 E_m}{\partial k_1\partial k_2}\right)^2+\frac{3}{8}\left(\frac{\partial E_m}{\partial k_1}\frac{\partial^3 E_m}{\partial k_1\partial k_2^2}+\frac{\partial E_m}{\partial k_2}\frac{\partial^3 E_m}{\partial k_1^2\partial k_2}\right)\right]\exp(-\beta E_m). \quad (45)$$

(b) *Reduction of* $Z^{(2)}$

We have to consider

$$Z^{(2)}=\frac{2}{(2\pi)^3}(\mu_0 H)^2\beta^2\sum_m\int d\boldsymbol{k}\left[-\frac{1}{2}\frac{\partial^2 E_m}{\partial k_2^2}+i\frac{\partial E_m}{\partial k_1}\frac{\partial Q_{mm}}{\partial k_2}+(Q_{mm})^2+\right.$$

$$+\frac{\partial E_m}{\partial k_1}\left(2iQ_{mm}-\frac{\partial^2 E_m}{\partial k_1\partial k_2}\right)U_{mm}-\left(\frac{\partial E_m}{\partial k_1}\right)^2 V_{mm}-\sum_{n\neq m}\frac{1}{E_n-E_m}\left\{\frac{\partial}{\partial k_2}\left(\frac{\partial E_m}{\partial k_2}P_{nm}\right)P_{mn}+\right.$$

$$\left.\left.+\frac{\partial E_m}{\partial k_2}(P_{mn}Q_{nm}+P_{nm}Q_{mn}+2P_{nm}P_{mn}U_{mm}+P_{mm}P_{mn}U_{mn}+3P_{mm}P_{nm}U_{nm})\right\}-\right.$$

$$\left.-\frac{\partial E_m}{\partial k_2}\sum_{n\neq m}\frac{1}{(E_n-E_m)^2}\left(2\frac{\partial E_m}{\partial k_2}-\frac{\partial E_n}{\partial k_2}\right)P_{nm}P_{mn}\right]\exp(-\beta E_m). \quad (46)$$

VOL. 13 1960

This is transformed as follows:

(i) The last two of the three terms involving U_{mm} are purely imaginary (note that $P_{nm} = -P_{mn}^*$), and give zero contribution.

(ii) We replace $\partial Q_{mm}/\partial k_2$ by its imaginary part, which is $\frac{1}{2}i\partial^3 E_m/\partial k_1 \partial k_2^2$, and $(Q_{mm})^2$ by its real part, which is

$$\left\{ik_1 \int u_m^* \frac{\partial u_m}{\partial k_2} d\tau_0 - \frac{1}{2}\int \left(\frac{\partial u_m^*}{\partial x}\frac{\partial u_m}{\partial k_2} + \frac{\partial u_m}{\partial x}\frac{\partial u_m^*}{\partial k_2}\right) d\tau_0\right\}^2 - \frac{1}{4}\left(\frac{\partial^2 E_m}{\partial k_1 \partial k_2}\right)^2.$$

(iii) We replace

$$(P_{mm}P_{mn}U_{mn} + 3P_{mm}P_{nm}U_{nm})/(E_n - E_m)$$

by

$$\{(P_{mm}P_{mn}U_{mn})^* + 3P_{mm}P_{nm}U_{nm}\}/(E_n - E_m)$$
$$= 2P_{mm}P_{nm}U_{nm}/(E_n - E_m)$$
$$= 2\frac{\partial E_m}{\partial k_1}\int u_n^* \frac{\partial u_n}{\partial k_1} d\tau_0 \int \psi_n \frac{\partial \psi_m^*}{\partial k_2} d\tau_0.$$

We do not need to carry out the summation over n for this term, since the term will be found to cancel later with a term obtained from $Z^{(1)} + Z^{(0)}$.

(iv) Using equations (A.3) and (A.6), we have

$$\sum_{n \neq m} \frac{1}{(E_n - E_m)^2} P_{nm}P_{mn} =$$
$$-\sum_{n \neq m} \int u_n^* \frac{\partial u_m}{\partial k_1} d\tau_0 \int u_n \frac{\partial u_m^*}{\partial k_1} d\tau_0$$
$$= -\int \left|\frac{\partial u_m}{\partial k_1}\right|^2 d\tau_0 + \left|\int u_m^* \frac{\partial u_m}{\partial k_1} d\tau_0\right|^2. \qquad (47)$$

(v) We replace

$$P_{mn}Q_{nm} + P_{nm}Q_{mn}$$

by

$$P_{mn}^* Q_{nm}^* + P_{nm}Q_{mn}$$
$$= P_{nm}(Q_{mn} - Q_{nm}^*)$$
$$= -P_{nm}\int \left(\frac{\partial u_m^*}{\partial x}\frac{\partial u_n}{\partial k_2} - \frac{\partial u_n}{\partial x}\frac{\partial u_m^*}{\partial k_2}\right) d\tau_0$$

(see equation (23)). An integration by parts allows this to be written as

$$P_{nm}\int \left(u_m^* \frac{\partial^2 u_n}{\partial x \partial k_2} + \frac{\partial u_m^*}{\partial k_2}\frac{\partial u_n}{\partial x}\right) d\tau_0$$
$$= P_{nm}\frac{\partial}{\partial k_2}\int u_m^* \frac{\partial u_n}{\partial x} d\tau_0$$
$$= P_{nm}\frac{\partial P_{mn}}{\partial k_2} \qquad (n \neq m).$$

Hence

$$\sum_{n \neq m} \frac{1}{E_n - E_m}\left\{\frac{\partial}{\partial k_2}\left(\frac{\partial E_m}{\partial k_2}P_{nm}\right)P_{mn} + \frac{\partial E_m}{\partial k_2}(P_{mn}Q_{nm} + P_{nm}Q_{mn})\right\}$$

becomes

$$\frac{\partial^2 E_m}{\partial k_2^2}\sum_{n \neq m}\frac{1}{E_n - E_m}P_{nm}P_{mn} + \frac{\partial E_m}{\partial k_2}\sum_{n \neq m}\frac{1}{E_n - E_m}\frac{\partial}{\partial k_2}(P_{nm}P_{mn})$$
$$= \left(\frac{\partial^2 E_m}{\partial k_2^2} + \frac{\partial E_m}{\partial k_2}\frac{\partial}{\partial k_2}\right)\sum_{n \neq m}\frac{1}{E_n - E_m}P_{nm}P_{mn} -$$
$$-\left(\frac{\partial E_m}{\partial k_2}\right)^2 \sum_{n \neq m}\frac{1}{(E_n - E_m)^2}P_{nm}P_{mn} + \frac{\partial E_m}{\partial k_2}\sum_{n \neq m}\frac{1}{(E_n - E_m)^2}\frac{\partial E_n}{\partial k_2}P_{nm}P_{mn}.$$

The last term cancels with the last term in equation (46), and the remaining terms may be evaluated by means of equations (42) and (47).

Using all these results, we find, after some simplification, that

$$Z^{(2)} = -\frac{2}{(2\pi)^3}(\mu_0 H)^2\beta^2 \sum_m \int dk \Big[\frac{1}{2}\frac{\partial^2 E_m}{\partial k_1^2}\frac{\partial^2 E_m}{\partial k_2^2}+\frac{1}{4}\Big(\frac{\partial^2 E_m}{\partial k_1 \partial k_2}\Big)^2+$$

$$+\frac{1}{2}\Big(\frac{\partial E_m}{\partial k_1}\frac{\partial^3 E_m}{\partial k_1 \partial k_2^2}+\frac{\partial E_m}{\partial k_2}\frac{\partial^3 E_m}{\partial k_1^2 \partial k_2}\Big)-2i\frac{\partial E_m}{\partial k_1}Q_{mm}U_{mm}+$$

$$+\Big(\frac{\partial E_m}{\partial k_1}\Big)^2 V_{mm}+2\frac{\partial E_m}{\partial k_1}\frac{\partial E_m}{\partial k_2}\sum_{n\neq m}\int u_n^*\frac{\partial u_m}{\partial k_1}\,d\tau_0\int \psi_n\frac{\partial \psi_m^*}{\partial k_2}\,d\tau_0-$$

$$-\Big\{ik_1\int u_m^*\frac{\partial u_m}{\partial k_2}\,d\tau_0-\frac{1}{2}\int\Big(\frac{\partial u_m^*}{\partial x}\frac{\partial u_m}{\partial k_2}+\frac{\partial u_m}{\partial x}\frac{\partial u_m^*}{\partial k_2}\Big)\,d\tau_0\Big\}^2-$$

$$-\Big(\frac{\partial E_m}{\partial k_2}\Big)^2\Big\{\int\Big|\frac{\partial u_m}{\partial k_1}\Big|^2 d\tau_0-\Big|\int u_m^*\frac{\partial u_m}{\partial k_1}\,d\tau_0\Big|^2\Big\}\Big]\exp(-\beta E_m). \qquad (48)$$

(c) *Reduction of* $Z^{(1)}+Z^{(0)}$

The remaining terms in Z may be simplified along similar lines, but the calculation is much more laborious and full details cannot be given here.

The first task is to deal with the double summations in $Z^{(0)}$ and $Z^{(1)}$. As an example consider the expression

$$\sum_{ln}^{l\neq n\neq m}\frac{1}{(E_l-E_m)(E_n-E_m)}\Big(P_{mn}\frac{\partial P_{nl}}{\partial k_2}U_{ml}+P_{lm}Q_{nl}U_{nm}\Big)$$

which appears among the terms in $Z^{(0)}$ (it consists of two of the terms obtained when equation (26) is substituted into the double summation in equation (36)). We shall show that one of the summations can be carried out in this expression. We interchange l and n in the second term and replace it by its complex conjugate. We then obtain, using equations (23), (A.3) and (A.6),

$$\sum_{ln}^{l\neq n\neq m}\frac{1}{(E_l-E_m)(E_n-E_m)}\Big(P_{mn}\frac{\partial P_{nl}}{\partial k_2}U_{ml}+P_{nm}^*Q_{ln}^*U_{lm}^*\Big)$$

$$=\sum_{ln}^{l\neq n\neq m}\frac{1}{(E_l-E_m)(E_n-E_m)}P_{mn}U_{ml}\Big(\frac{\partial P_{nl}}{\partial k_2}+Q_{ln}^*\Big)$$

$$=i\sum_{ln}^{l\neq n\neq m}\frac{U_{ml}}{E_l-E_m}\int u_m^*\frac{\partial u_n}{\partial k_1}\,d\tau_0\Big\{\frac{\partial}{\partial k_2}\int u_n^*\frac{\partial u_l}{\partial x}\,d\tau_0-$$

$$-ik_1\int u_l\frac{\partial u_n^*}{\partial k_2}\,d\tau_0-\int\frac{\partial u_l}{\partial x}\frac{\partial u_n^*}{\partial k_2}\,d\tau_0\Big\}$$

$$=-i\sum_{ln}^{l\neq n\neq m}\frac{U_{ml}}{E_l-E_m}\int u_n\frac{\partial u_m^*}{\partial k_1}\,d\tau_0\Big(\int u_n^*\frac{\partial^2 u_l}{\partial x\partial k_2}\,d\tau_0+ik_1\int u_n^*\frac{\partial u_l}{\partial k_2}\,d\tau_0\Big).$$

This is now in a form in which the summation over n can be carried out at once (the result obtained will not be written out here). It is found that, by similar considerations, at least one summation can be carried out in every double sum that appears in $Z^{(0)}$ and $Z^{(1)}$.

We next collect together and simplify the terms in $Z^{(0)}$. By means of a series of integrations by parts of the general types

$$\int\frac{\partial F}{\partial k_1}\exp(-\beta E_m)\,d\boldsymbol{k}$$

$$=\beta\int\frac{\partial E_m}{\partial k_1}F\exp(-\beta E_m)\,d\boldsymbol{k}, \tag{49}$$

$$\int\frac{\partial F}{\partial k_1}G\exp(-\beta E_m)\,d\boldsymbol{k}$$

$$=\beta\int\frac{\partial E_m}{\partial k_1}FG\exp(-\beta E_m)\,d\boldsymbol{k}-$$

$$-\int F\frac{\partial G}{\partial k_1}\exp(-\beta E_m)\,d\boldsymbol{k}, \tag{50}$$

etc. (the surface integrals all vanish), the complete expression for $Z^{(0)}$ can eventually be transformed into a form in which the integrand depends on β only through a factor $\beta\exp(-\beta E_m)$, the "zero-order" terms cancelling exactly. The transformed expression for $Z^{(0)}$ is then combined with $Z^{(1)}$, and the result is simplified again by means of transformations of the type (49) and (50). When all possible reductions have been made, we finally obtain $Z^{(1)}+Z^{(0)}$ in a form that contains both "linear" terms whose integrands involve the factor $\beta\exp(-\beta E_m)$, and "quadratic" terms whose integrands involve the factor $\beta^2\exp(-\beta E_m)$ and which can be combined with equations (45) and (48). After lengthy algebra the surviving terms are found to be

$$Z^{(1)}+Z^{(0)}=-\frac{2}{(2\pi)^3}(\mu_0H)^2\beta^2\sum_m\int d\boldsymbol{k}\Big[2i\frac{\partial E_m}{\partial k_1}Q_{mm}U_{mm}-\Big(\frac{\partial E_m}{\partial k_1}\Big)^2V_{mm}+$$

$$+2i\frac{\partial E_m}{\partial k_1}\frac{\partial E_m}{\partial k_2}\sum_{n\neq m}X_{nm}\int\psi_n\frac{\partial\psi_m^*}{\partial k_2}\,d\tau_0+2\Big(\frac{\partial E_m}{\partial k_2}\Big)^2\Big(\int\Big|\frac{\partial u_m}{\partial k_1}\Big|^2d\tau_0-|X_{mm}|^2\Big)+$$

$$+\frac{\partial E_m}{\partial k_2}\Big\{-k_1\int\Big(\frac{\partial u_m^*}{\partial k_1}\frac{\partial u_m}{\partial k_2}+\frac{\partial u_m}{\partial k_1}\frac{\partial u_m^*}{\partial k_2}\Big)\,d\tau_0+2X_{mm}W_m+$$

$$+i\int\Big(\frac{\partial u_m^*}{\partial k_1}\frac{\partial^2u_m}{\partial x\partial k_2}-\frac{\partial u_m}{\partial k_1}\frac{\partial^2u_m^*}{\partial x\partial k_2}\Big)\,d\tau_0\Big\}\Big]\exp(-\beta E_m)-$$

$$-\frac{2}{(2\pi)^3}(\mu_0H)^2\beta\sum_m\int d\boldsymbol{k}\Big[\int\Big|\frac{\partial u_m}{\partial k_2}\Big|^2d\tau_0-2\sum_{n\neq m}\frac{1}{E_n-E_m}\Big|k_1Y_{nm}-\int\frac{\partial u_n^*}{\partial x}\frac{\partial u_m}{\partial k_2}\,d\tau_0\Big|^2+$$

$$+2\frac{\partial E_m}{\partial k_2}\sum_{n\neq m}\frac{X_{nm}}{E_n-E_m}\Big\{\int\Big(\frac{\partial u_n}{\partial y}\frac{\partial u_m^*}{\partial k_1}-\frac{\partial u_n}{\partial x}\frac{\partial u_m^*}{\partial k_2}\Big)\,d\tau_0+$$

$$+k_1 Y_{mn}-k_2 X_{mn}+\frac{\partial E_m}{\partial k_1}Y_{mn}\Big\}+$$

$$+i\frac{\partial E_m}{\partial k_2}X_{mm}\int\left(\frac{\partial u_m^*}{\partial k_1}\frac{\partial u_m}{\partial k_2}-\frac{\partial u_m}{\partial k_1}\frac{\partial u_m^*}{\partial k_2}\right)d\tau_0+2W_m\frac{\partial X_{mm}}{\partial k_2}\Big]\exp(-\beta E_m), \quad (51)$$

where
$$X_{mn}=i\int u_m^*\frac{\partial u_n}{\partial k_1}d\tau_0, \quad Y_{mn}=i\int u_m^*\frac{\partial u_n}{\partial k_2}d\tau_0, \quad (52)$$

and
$$W_m=k_1 Y_{mm}-\frac{1}{2}\int\left(\frac{\partial u_m^*}{\partial x}\frac{\partial u_m}{\partial k_2}+\frac{\partial u_m}{\partial x}\frac{\partial u_m^*}{\partial k_2}\right)d\tau_0.$$

The complete expression for Z (to order H^2) is obtained by adding equations (45), (48) and (51). It is seen that the first three terms in (51) cancel with terms in (48), while the fourth term in (51) changes the sign of the last term in (48).

5. THE DIAMAGNETIC SUSCEPTIBILITY

The expression for Z will not be written out again, and we proceed at once to transform to the general case of Fermi statistics. The term proportional to H^2 in the Fermi free energy F may be written down at once[3] by replacing a factor $\beta\exp(-\beta E_m)$ in Z by $-2f_0(E_m)$, and a factor $\beta^2\exp(-\beta E_m)$ by $2\partial f_0(E_m)/\partial E_m$, where $f_0(E)$ is the Fermi function. The volume susceptibility χ is then obtained as $-(1/H)\partial F/\partial H$ (the H-dependence of the Fermi level may be neglected in calculating the constant term in the susceptibility). The final expression for the steady susceptibility of a crystal of arbitrary symmetry, for a magnetic field applied in the z-direction, is conveniently written as the sum of four terms, $\chi=\chi_1+\chi_2+\chi_3+\chi_4$, where, on changing from atomic to arbitrary units, we obtain†

$$\chi_1=\frac{e^2}{48\pi^3\hbar^2c^2}\sum_m\int\left\{\frac{\partial^2E_m}{\partial k_1^2}\frac{\partial^2E_m}{\partial k_2^2}-\left(\frac{\partial^2E_m}{\partial k_1\partial k_2}\right)^2+\frac{3}{2}\left(\frac{\partial E_m}{\partial k_1}\frac{\partial^3E_m}{\partial k_1\partial k_2^2}+\frac{\partial E_m}{\partial k_2}\frac{\partial^3E_m}{\partial k_1^2\partial k_2}\right)\right\}\frac{\partial f_0(E_m)}{\partial E_m}d\boldsymbol{k}, \quad (53)$$

$$\chi_2=\frac{e^2}{4\pi^3mc^2}\sum_m\int\Big[\frac{m}{\hbar^2}\frac{\partial E_m}{\partial k_2}\Big\{\int\left|\frac{\partial u_m}{\partial k_1}\right|^2d\tau_0-|X_{mm}|^2\Big\}-$$

$$-k_1\int\left(\frac{\partial u_m^*}{\partial k_1}\frac{\partial u_m}{\partial k_2}+\frac{\partial u_m}{\partial k_1}\frac{\partial u_m^*}{\partial k_2}\right)d\tau_0+2X_{mm}W_m+$$

$$+i\int\left(\frac{\partial u_m^*}{\partial k_1}\frac{\partial^2u_m}{\partial x\partial k_2}-\frac{\partial u_m}{\partial k_1}\frac{\partial^2u_m^*}{\partial x\partial k_2}\right)d\tau_0\Big]\frac{\partial E_m}{\partial k_2}\frac{\partial f_0(E_m)}{\partial E_m}d\boldsymbol{k}, \quad (54)$$

$$\chi_3=-\frac{e^2}{4\pi^3mc^2}\sum_m\int\Big[\frac{im}{\hbar^2}X_{mm}\int\left(\frac{\partial u_m^*}{\partial k_1}\frac{\partial u_m}{\partial k_2}-\frac{\partial u_m}{\partial k_1}\frac{\partial u_m^*}{\partial k_2}\right)d\tau_0+$$

$$+2\sum_{n\neq m}\frac{X_{nm}}{E_n-E_m}\Big\{\int\left(\frac{\partial u_n}{\partial y}\frac{\partial u_m^*}{\partial k_1}-\frac{\partial u_n}{\partial x}\frac{\partial u_m^*}{\partial k_2}\right)d\tau_0+$$

$$+k_1Y_{mn}-k_2X_{mn}+\frac{m}{\hbar^2}\frac{\partial E_m}{\partial k_1}Y_{mn}\Big\}\Big]\frac{\partial E_m}{\partial k_2}f_0(E_m)\,d\boldsymbol{k}, \quad (55)$$

† Note that the present notation differs slightly from that used by HEBBORN and SONDHEIMER[9]. The use of m in two different senses, as a band index and to denote the free electron mass, should not cause any confusion.

$$\chi_4 = -\frac{e^2}{4\pi^3 mc^2}\sum_m \int\Bigg[\Bigg\{\int\left|\frac{\partial u_m}{\partial k_2}\right|^2 d\tau_0 + 2W_m\frac{\partial X_{mm}}{\partial k_2} -$$

$$-2\frac{\hbar^2}{m}\sum_{n\neq m}\frac{1}{E_n-E_m}\left|k_1 Y_{nm} - \int\frac{\partial u_n^*}{\partial x}\frac{\partial u_m}{\partial k_2}d\tau_0\right|^2\Bigg\} f_0(E_m) + \frac{\hbar^2}{m}(W_m)^2\frac{\partial f_0(E_m)}{\partial E_m}\Bigg] d\boldsymbol{k}. \quad (56)$$

We recall that X_{mn}, Y_{mn} and W_m are defined by equation (52).

Obviously many equivalent forms of the final result are possible, but we have not been able to reduce the formulae for the general case to a form that is significantly simpler than that given here. The structure of the terms in the general expression for χ is considered further below.

(a) *Invariance properties*

It is important to verify, as a check on the calculation, that the expression for χ satisfies certain general invariance requirements. We note, first, that the result must be independent of the choice of phase for the Bloch functions,† and must therefore be unchanged in form under the transformation

$$u_{\boldsymbol{k}m} = u'_{\boldsymbol{k}m}\exp\{i\theta_m(\boldsymbol{k})\}, \quad (57)$$

where time-reversal invariance requires that

$$\theta_m(-\boldsymbol{k}) = -\theta_m(\boldsymbol{k}).$$

It is easy to see that, for example, the factor

$$\int\left|\frac{\partial u_m}{\partial k_1}\right|^2 d\tau_0 - |X_{mm}|^2$$

in χ_2 is invariant under this transformation, and in fact it can be shown that each of the four terms in χ is separately invariant by explicit verification that the additional terms, involving the derivatives of $\theta_m(\boldsymbol{k})$, which arise when equation (57) is substituted into (53) to (56), are identically zero.

For a crystal that possesses a centre of symmetry at $\boldsymbol{r} = 0$, we may simplify the general expression for χ by choosing the phase of the Bloch functions such that $u_{\boldsymbol{k}m}(-\boldsymbol{r}) = u^*_{\boldsymbol{k}m}(\boldsymbol{r})$. We then have

$$\int u_m^*\frac{\partial u_n}{\partial k_1}d\tau_0 = \int u_m\frac{\partial u_n^*}{\partial k_1}d\tau_0,$$

$$\int\frac{\partial u_m^*}{\partial x}\frac{\partial u_m}{\partial k_2}d\tau_0 = -\int\frac{\partial u_m}{\partial x}\frac{\partial u_m^*}{\partial k_2}d\tau_0$$

(and similar relations), so that X_{mn} and Y_{mn} are purely imaginary, and $X_{mm} = Y_{mm} = W_m = 0$. In this case equations (53) to (56) reduce to the somewhat simpler expressions published earlier.[9] For cubic lattices, moreover, χ_3 reduces to the single term

$$\chi_3 = -\frac{e^2}{2\pi^3\hbar^2c^2}\sum_m\int\Bigg\{\frac{\partial E_m}{\partial k_1}\frac{\partial E_m}{\partial k_2}\times$$

$$\times\sum_{n\neq m}\frac{1}{E_n-E_m}\int u_m^*\frac{\partial u_n}{\partial k_1}d\tau_0\times$$

$$\times\int u_m^*\frac{\partial u_n}{\partial k_2}d\tau_0\Bigg\} f_0(E_m)\,d\boldsymbol{k}. \quad (58)$$

We must also consider gauge invariance. The particular algebraic form taken by our final expression for χ is determined by the choice $(-Hy, 0, 0)$ for the vector potential that corresponds to the magnetic field $(0, 0, H)$. We could equally well have chosen the vector potential in the form $(0, Hx, 0)$, and would then have obtained a result in which x and y, and k_1 and k_2, are interchanged wherever they appear explicitly in the formulae (53) to (56). And, by taking half the sum of this expression and the original one, we can write χ in a form that is symmetrical in x- and y-components and evidently corresponds (in some sense) to the vector potential $(-\frac{1}{2}Hy, \frac{1}{2}Hx, 0)$. These symmetry properties can be demonstrated by direct transformation of equations (53) to (56).

† The importance of this requirement has been pointed out to us by Professor W. KOHN. We are also indebted to Professor HARVEY BROOKS for helpful comments on this matter and on the interpretation of χ.

For example, consider the terms (taken from χ_2 and χ_4)

$$\left(\frac{\partial E_m}{\partial k_2}\right)^2 |X_{mm}|^2 - 2\frac{\partial E_m}{\partial k_2} X_{mm} W_m + W_m^2 = \left(\frac{\partial E_m}{\partial k_2} X_{mm} - W_m\right)^2$$

$$= \left\{ i\frac{\partial E_m}{\partial k_2} \int u_m^* \frac{\partial u_m}{\partial k_1} d\tau_0 - ik_1 \int u_m^* \frac{\partial u_m}{\partial k_2} d\tau_0 + \frac{1}{2}\int \left(\frac{\partial u_m^*}{\partial x}\frac{\partial u_m}{\partial k_2} + \frac{\partial u_m}{\partial x}\frac{\partial u_m^*}{\partial k_2} \right) d\tau_0 \right\}^2. \quad (59)$$

(we revert to atomic units and recall that X_{mm} is real). In Appendix 2 we show that this equals

$$\left\{ i\frac{\partial E_m}{\partial k_1} \int u_m^* \frac{\partial u_m}{\partial k_2} d\tau_0 - ik_2 \int u_m^* \frac{\partial u_m}{\partial k_1} d\tau_0 + \right.$$

$$\left. + \frac{1}{2}\int \left(\frac{\partial u_m^*}{\partial y}\frac{\partial u_m}{\partial k_1} + \frac{\partial u_m}{\partial y}\frac{\partial u_m^*}{\partial k_1} \right) d\tau_0 \right\}^2 \quad (60)$$

and thus possesses the required symmetry.

The complete proofs of these invariance properties involve only transformations of the same type as were used in the calculations of Section 4. Since the calculations are somewhat lengthy and do not possess any novel features, the details will be omitted here.

(b) *Special cases*

(i) *Free electrons.* For perfectly free electrons $u_{\boldsymbol{k}}$ is constant and $E_{\boldsymbol{k}}$ is proportional to $|\boldsymbol{k}|^2$ (there is only one "band"). It follows immediately that the only surviving term in χ is the part of χ_1 that involves the second derivatives of the energy. This is just the Peierls term χ_P which gives the correct free-electron diamagnetism.

(ii) *Tightly bound electrons.* In the tight-binding limit of large separation of the atoms, all the energy derivatives tend to zero exponentially with increasing lattice parameter, and only χ_4 survives amongst the terms in χ. If we neglect all exponentially vanishing terms, we may write

$$u_{\boldsymbol{k}m}(\boldsymbol{r}) = \exp(-i\boldsymbol{k}\cdot\boldsymbol{r})\phi_m(\boldsymbol{r}), \quad (61)$$

where $\phi_m(\boldsymbol{r})$, which is taken to be real, is an atomic function within the lattice cell centred at $\boldsymbol{r} = 0$ (whose radius tends to infinity). Thus

$$\frac{\partial u_m}{\partial k_2} = -iy\phi_m \exp(-i\boldsymbol{k}\cdot\boldsymbol{r}),$$

$$\frac{\partial u_m}{\partial x} = \left(\frac{\partial \phi_m}{\partial x} - ik_1\phi_m\right)\exp(-i\boldsymbol{k}\cdot\boldsymbol{r}),$$

so that $W_m = 0$† and we obtain for the contribution $\chi_{4,m}$ to χ_4 from the atomic state m (the integration in (56) over the occupied region in $\boldsymbol{k}$-space gives a factor $\frac{1}{2}(2\pi)^3 n$, where n is the number of conduction electrons per unit volume)

$$\chi_{4,m} = -\frac{ne^2}{mc^2}\left\{ \int y^2\phi_m^2 \, d\boldsymbol{r} - \right.$$

$$\left. -2\frac{\hbar^2}{m}\sum_{n\neq m}\frac{1}{E_n - E_m}\left| \int \phi_m y \frac{\partial \phi_n}{\partial x} d\boldsymbol{r} \right|^2 \right\}. \quad (62)$$

This is just the usual formula (expressed in the unsymmetrical gauge used here) for the ordinary atomic diamagnetism, including the "VAN VLECK paramagnetism".[12] In particular, when the state $\boldsymbol{m}$ is an s-state, the summation in equation (62) may be evaluated and gives

$$\sum_{n\neq m}\frac{1}{E_n - E_m}\left| \int \phi_m y \frac{\partial \phi_n}{\partial x} d\boldsymbol{r} \right|^2 = \frac{m}{4\hbar^2}\int y^2\phi_m^2 \, d\boldsymbol{r}.$$

In this case (62) reduces to the familiar expression

$$\chi_{4,m} = -\frac{ne^2}{6mc^2}(r^2)_{mm}. \quad (63)$$

(iii) *Nearly empty band.* The special case of a small number of degenerate electrons, contained near a band minimum (at $\boldsymbol{k} = 0$) in a cubic crystal, has been considered by KJELDAAS and KOHN[6]. These authors calculate χ as a power series in ascending powers of k_0, where $k_0^3 = 3\pi^2 n$, of the form

$$\chi = ak_0 + bk_0^3 + \ldots, \quad (64)$$

in which the leading term ak_0 arises entirely from the Peierls term χ_P, whereas the term bk_0^3 contains,

† W_m is in fact proportional to the magnetic moment of the atomic state m and, when this is non-zero, the last term in χ_4 leads to a non-zero Langevin–Debye paramagnetic susceptibility.

VOL. 13 1960

in addition to a contribution from χ_P, a term χ_R which is shown by KJELDAAS and KOHN to reduce to the atomic diamagnetism in the tight-binding limit.

It can be shown that, in the approximation represented by equation (64), the result (53) to (56) given by the present theory reduces to that obtained by KJELDAAS and KOHN. To obtain the analogue of the KJELDAAS–KOHN expansion, we expand u_{km} and E_{km} about $\boldsymbol{k} = 0$ in powers of the components of $\boldsymbol{k}$. It is then clear that, since all the non-zero terms in χ except χ_P can be written as volume integrals over the occupied regions in $\boldsymbol{k}$-space,* their entire contribution to the expansion (64) must be of the order of k_0^3, and we therefore conclude, in agreement with KJELDAAS and KOHN, that when terms of order k_0^3 are neglected the result given by χ_P is exact. It is a more difficult matter to show that the terms of order k_0^3 are equivalent in the two theories, and, in view of the restricted range of validity of the result, the details of the proof are omitted here.

(c) *Discussion*

The general expressions (53) to (56) are valid for arbitrary degeneracy of the electron gas. (We note in passing that, while χ is independent of temperature for highly degenerate electrons, and while its temperature variation for arbitrary degeneracy is easily calculated for perfectly free electrons (WILSON[10], p. 167), it is not possible to derive any simple law for the temperature variation of χ for the general case of arbitrary Fermi–Dirac degeneracy and arbitrary E–$\boldsymbol{k}$ relations. A detailed investigation of the temperature variation to be expected for particular models would be of interest.) In the subsequent discussion we shall, with applications to metals in mind, confine ourselves to the case of high degeneracy. In this case χ_1 and χ_2 may be expressed as surface integrals over the Fermi surface (which need not lie in a single band), while χ_3 is a volume integral extending over the occupied regions in every band and χ_4 is a mixed volume and surface integral (but the separation into surface and volume integrals in $\boldsymbol{k}$-space is not unique).

The term χ_1 represents that part of χ that can be expressed in terms of the energy derivatives only, and that is therefore determined entirely by the energy level structure in the vicinity of the Fermi surface. It seems to have been generally assumed hitherto that the only term in χ that could be expressed in this way was the Peierls term χ_P, but we now see that χ_1 contains, in addition to χ_P, a second term (χ_Q, say) that involves the third derivatives of the energy and that does not seem to possess any simple physical interpretation. For a nearly full or a nearly empty band, where $E_{\boldsymbol{k}}$ is approximately quadratic, χ_P is negative and large compared with χ_Q, but for more general E–$\boldsymbol{k}$ relations χ_P and χ_Q will generally be of comparable magnitude, and each may be either negative or positive. (When E is proportional to $|\boldsymbol{k}|^4$, for example, $\chi_P/\chi_Q = 7/6$. But there is little point in comparing χ_P and χ_Q for particular models without also considering the contributions from χ_2, χ_3 and χ_4. In fact the division of χ into the sum of separate terms, and in particular the contribution assigned to χ_1, is to some extent arbitrary and may be affected, for example, by the choice of gauge; it is this circumstance that leads to the appearance of so many superficially quite dissimilar expressions for χ in the literature.)

Of particular interest are the highly diamagnetic metals such as bismuth and the γ-phase copper–zinc alloys, where the susceptibility is supposed to arise from small numbers of electrons, of small effective mass, associated with high-curvature portions of the Fermi surface lying close to a zone boundary. If we suppose as a crude approximation that the energy spectrum is that of quasi-free electrons, with $E = \hbar^2|\boldsymbol{k}|^2/2\boldsymbol{m}^*$, we have $\chi_Q = 0$, but χ_P is $\boldsymbol{m}/\boldsymbol{m}^*$ times as large as the free-electron value (for a given carrier density) and is usually regarded as giving the dominant contribution to χ. This assumption has been criticized by ADAMS[5], who pointed out that, when $\boldsymbol{m}^*$ is small, there must be other terms in χ of abnormally large magnitude, and it is easy to see qualitatively that the present theory leads to similar conclusions. For example, when we compare χ_2 and χ_3 with the analogue χ_4 of the atomic diamagnetism, we easily see that the first term in χ_2 is of order $\boldsymbol{m}/\boldsymbol{m}^*$ times the first term in χ_4, while χ_3 contains a term of order $(\boldsymbol{m}/\boldsymbol{m}^*)^2$ times

* For example, the second part of χ_1 can be written in the form

$$-\frac{e^2}{16\pi^3\hbar^2c^2}\sum_m\int\frac{\partial^4 E_m}{\partial k_1^2\partial k_2^2}f_0(E_m)\,d\boldsymbol{k}.$$

Note that $X_{mm} = W_m = 0$ for a cubic lattice.

a corresponding term in χ_4; and therefore, even when χ_P is large compared with free-electron and atomic diamagnetic susceptibilities, it is not necessarily large compared with the "mixed" terms χ_2 and χ_3.

Considerations such as these are, of course, much too tentative to be in any way conclusive, and with equations (53) to (56) we are still far from having numerical values of χ for particular substances. But at least we now have a complete formula (subject to the basic assumptions made in the theory) which allows χ to be calculated, for any metal or semiconductor, as soon as sufficiently detailed information is available about the Bloch wave functions and energy levels. We shall not discuss here the problems involved in the computational aspects of the subject, nor shall we enter into the practical problem of obtaining the diamagnetic contribution of the conduction electrons from experimentally measured susceptibilities. We remark only that, since the formula for χ involves the second derivatives of u_k and the third derivatives of E_k, the calculated values of χ are likely to be sensitive to the finer details of the energy-level spectrum and the precise form of the variation of u_k with r and k, except in simple cases (if any exist) where the free-electron approximation is not too far from the truth. Also the fact that χ_3 and χ_4 still contain non-zero contributions from summations over excited bands will, in general, seriously complicate the numerical analysis, unless these contributions can be shown to be small (but one must here remember the sum rules and the fact that the identification of the terms to be regarded as arising from excited bands depends upon the gauge of the vector potential). Even when all simplifications arising from crystal symmetry have been made, the computational obstacles are therefore likely to prove formidable for any but the simpler monovalent metals, although it would perhaps be unwise to over-emphasize the difficulties at this stage.

Acknowledgements—We are indebted to many colleagues for helpful discussions of various aspects of this problem. Part of the work was carried out while one of the authors (E.H.S.) was visiting the Physics Departments of the Universities of Illinois and California (Berkeley). He wishes to express his thanks to Professor F. SEITZ and Professor C. KITTEL and the staffs of these Departments for their generous hospitality. The other author (J.E.H. wishes to thank the Department of Scientific and Industrial Research for the award of a Research Studentship.

APPENDIX 1

Properties of Bloch Functions

In the reduction of the multiple integrals in Section 2 we require the "Fourier formulae" (JONES and ZENER[13]; WILSON[10], Section 2.82):

$$\frac{\Delta}{(2\pi)^3}\int d\boldsymbol{r}\int d\boldsymbol{k}'F(\boldsymbol{k}')\psi^*_{\boldsymbol{k}m}(\boldsymbol{r})y\frac{\partial}{\partial x}\psi_{\boldsymbol{k}'n}(\boldsymbol{r})$$

$$= i\frac{\partial F(\boldsymbol{k})}{\partial k_2}\int \psi^*_{\boldsymbol{k}m}\frac{\partial \psi_{\boldsymbol{k}n}}{\partial x}d\tau_0 -$$

$$-F(\boldsymbol{k})\Big\{k_1\int u^*_{\boldsymbol{k}m}\frac{\partial u_{\boldsymbol{k}n}}{\partial k_2}d\tau_0 +$$

$$+i\int \frac{\partial u^*_{\boldsymbol{k}m}}{\partial x}\frac{\partial u_{\boldsymbol{k}n}}{\partial k_2}d\tau_0\Big\}, \quad \text{(A.1)}$$

$$\frac{\Delta}{(2\pi)^3}\int d\boldsymbol{r}\int d\boldsymbol{k}'F(\boldsymbol{k}')\psi^*_{\boldsymbol{k}m}(\boldsymbol{r})y^2\psi_{\boldsymbol{k}'n}(\boldsymbol{r})$$

$$= -\frac{\partial^2 F(\boldsymbol{k})}{\partial k_2^2}\delta_{mn} - 2\frac{\partial F(\boldsymbol{k})}{\partial k_2}\int u^*_{\boldsymbol{k}m}\frac{\partial u_{\boldsymbol{k}n}}{\partial k_2}d\tau_0 -$$

$$-F(\boldsymbol{k})\int u^*_{\boldsymbol{k}m}\frac{\partial^2 u_{\boldsymbol{k}n}}{\partial k_2^2}d\tau_0, \quad \text{(A.2)}$$

in which $F(\boldsymbol{k})$ is any twice differentiable periodic function of the reduced wave vector.

For fixed $\boldsymbol{k}$, we have the following well-known relations among matrix elements of Bloch functions. (See WILSON[10], Chap. 2. We recall that atomic units are used here, and omit $\boldsymbol{k}$ where no ambiguity is possible).

$$P_{mn} \equiv \int \psi^*_m\frac{\partial \psi_n}{\partial x}d\tau_0 = ik_1\delta_{mn} +$$

$$+\int u^*_m\frac{\partial u_n}{\partial x}d\tau_0 = i\frac{\partial E_m}{\partial k_1}\delta_{mn} +$$

$$+i(E_n - E_m)\int u^*_m\frac{\partial u_n}{\partial k_1}d\tau_0, \quad \text{(A.3)}$$

$$i\int\left(\frac{\partial u_m}{\partial x}\frac{\partial u^*_m}{\partial k_1} - \frac{\partial u^*_m}{\partial x}\frac{\partial u_m}{\partial k_1}\right)d\tau_0 = 1 - \frac{\partial^2 E_m}{\partial k_1^2}, \quad \text{(A.4)}$$

VOL. 13 1960

(with corresponding expression in which x, k_1 are replaced by y, k_2), and

$$i\int\left(\frac{\partial u_m}{\partial x}\frac{\partial u_m^*}{\partial k_2}-\frac{\partial u_m^*}{\partial x}\frac{\partial u_m}{\partial k_2}\right)d\tau_0$$

$$=i\int\left(\frac{\partial u_m}{\partial y}\frac{\partial u_m^*}{\partial k_1}-\frac{\partial u_m^*}{\partial y}\frac{\partial u_m}{\partial k_1}\right)d\tau_0$$

$$=-\frac{\partial^2 E_m}{\partial k_1\partial k_2}. \qquad \text{(A.5)}$$

We note also that

$$\int\psi_m^*\frac{\partial\psi_n}{\partial k_2}d\tau_0=-\int\psi_n\frac{\partial\psi_m^*}{\partial k_2}d\tau_0,$$

$$\int u_m^*\frac{\partial u_n}{\partial k_2}d\tau_0=-\int u_n\frac{\partial u_m^*}{\partial k_2}d\tau_0, \qquad \text{(A.6)}$$

so that

$$\int\psi_m^*\frac{\partial\psi_m}{\partial k_2}d\tau_0 \text{ and } \int u_m^*\frac{\partial u_m}{\partial k_2}d\tau_0$$

are pure imaginary quantities.

APPENDIX 2

Equivalence of Equations (59) *and* (60)

We have, using familiar transformations,

$$\int\frac{\partial u_m^*}{\partial x}\frac{\partial u_m}{\partial k_2}d\tau_0+i\left(\frac{\partial E_m}{\partial k_1}-k_1\right)\int u_m^*\frac{\partial u_m}{\partial k_2}d\tau_0$$

$$=\int\frac{\partial u_m^*}{\partial x}\frac{\partial u_m}{\partial k_2}d\tau_0+\int u_m^*\frac{\partial u_m}{\partial x}d\tau_0\int u_m^*\frac{\partial u_m}{\partial k_2}d\tau_0$$

$$=\sum_{n\neq m}\int u_n\frac{\partial u_m^*}{\partial x}d\tau_0\int u_n^*\frac{\partial u_m}{\partial k_2}d\tau_0$$

$$=i\sum_{n\neq m}(E_n-E_m)\int u_n\frac{\partial u_m^*}{\partial k_1}d\tau_0\int u_n^*\frac{\partial u_m}{\partial k_2}d\tau_0$$

$$=-\sum_{n\neq m}\int u_n\frac{\partial u_m^*}{\partial k_1}d\tau_0\int u_n^*\frac{\partial u_m}{\partial y}d\tau_0$$

$$=-\int\frac{\partial u_m}{\partial y}\frac{\partial u_m^*}{\partial k_1}d\tau_0-\int u_m^*\frac{\partial u_m}{\partial y}d\tau_0\int u_m^*\frac{\partial u_m}{\partial k_1}d\tau_0$$

$$=-\int\frac{\partial u_m}{\partial y}\frac{\partial u_m^*}{\partial k_1}d\tau_0-i\left(\frac{\partial E_m}{\partial k_2}-k_2\right)\int u_m^*\frac{\partial u_m}{\partial k_1}d\tau_0.$$

Hence, rearranging,

$$i\frac{\partial E_m}{\partial k_2}\int u_m^*\frac{\partial u_m}{\partial k_1}d\tau_0-ik_1\int u_m^*\frac{\partial u_m}{\partial k_2}d\tau_0+\int\frac{\partial u_m^*}{\partial x}\frac{\partial u_m}{\partial k_2}d\tau_0$$

$$=-i\frac{\partial E_m}{\partial k_1}\int u_m^*\frac{\partial u_m}{\partial k_2}d\tau_0+ik_2\int u_m^*\frac{\partial u_m}{\partial k_1}d\tau_0-\int\frac{\partial u_m}{\partial y}\frac{\partial u_m^*}{\partial k_1}d\tau_0,$$

from which the equivalence of equations (59) and (60) follows on taking real parts.

REFERENCES

1. LANDAU L., *Z. Phys.* **64,** 629 (1930).
2. KOHN W., *Proc. Phys. Soc., Lond.* **72,** 1147 (1958).
3. WILSON A. H., *Proc. Camb. Phil. Soc.* **49,** 292 (1953).
4. PEIERLS R., *Z. Phys.* **80,** 763 (1933).
5. ADAMS E. N., *Phys. Rev.* **89,** 633 (1953).
6. KJELDAAS T. and KOHN W., *Phys. Rev.* **105,** 806 (1957).
7. WENTZEL G., *Phys. Rev.* **108,** 1593 (1957).
8. SONDHEIMER E. H. and WILSON A. H., *Proc. Roy. Soc.* A**210,** 173 (1951).
9. HEBBORN J. E. and SONDHEIMER E. H., *Phys. Rev. Letters* **2,** 150 (1959).
10. WILSON A. H., *The Theory of Metals* (2nd ed.). Cambridge University Press (1953).
11. GOLDBERGER M. L. and ADAMS E. N., *J. Chem. Phys.* **20,** 240 (1952).
12. VAN VLECK J. H., *The Theory of Electric and Magnetic Susceptibilities* p. 227. Clarendon Press, Oxford (1932).
13. JONES H. and ZENER C., *Proc. Roy. Soc.* A**144,** 101 (1934).

VOL.
13
1960

Note added in proof: An apparently simpler method has recently been used by ENZ [ENZ C. P., *Nuovo Cim.* 6 (Suppl.), 1224 (1957); *Helv. Phys. Acta* (to be published)] to treat the problem considered in the present paper, but the final result differs from that obtained here. The discrepancy arises from ENZ's use of an incorrect definition of the trace of an operator when the normalization volume is infinite. We thank Dr. ENZ for sending us a pre-publication copy of his paper and for discussions of the comparison of the two theories.

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 13. pp. 124–131. Printed in Great Britain.

CORRELATION OF THERMOLUMINESCENCE IN MgO AND VALENCE CHANGES OF IRON AND CHROMIUM IMPURITIES DETECTED BY EPR

R. L. HANSLER and W. G. SEGELKEN

Lamp Research Laboratory, General Electric Company, Nela Park, Cleveland 12, Ohio

(*Received* 9 *October* 1959)

Abstract—EPR measurements show that during 2537 Å irradiation the valence of chromium and iron impurities (~10 p.p.m.) in MgO single crystals changes from 2 to 3 or vice versa dependent on treatment. The reverse valence changes as well as thermoluminescence are observed during subsequent heating. It was shown by optical absorption that the energy levels of Fe^{+3} and Cr^{+3} lie about 5·8 and 6·0 eV respectively above the valence band while the levels of Fe^{+2} and Cr^{+2} are respectively 6·9 and 5·1 below the conduction band. The possible luminescent processes are:

$$Fe^{+3} + \text{free electron} \rightarrow Fe^{+2} + 2{\cdot}9\,\text{eV};$$
$$Fe^{+2} + \text{free hole} \rightarrow Fe^{+3} + 2{\cdot}0\,\text{eV};$$
$$Cr^{+3} + \text{free electron} \rightarrow Cr^{+2} + 2{\cdot}7\,\text{eV};$$
$$Cr^{+2} + \text{free hole} \rightarrow Cr^{+3} + 3{\cdot}6\,\text{eV}.$$

Four crystals were studied each of which showed preponderantly only one of the above luminescent processes. By assuming the band gap is 8·7 eV, good agreement between the color of the observed thermoluminescence and that predicted by the model was obtained.

1. INTRODUCTION

THE luminescence of MgO when excited in various ways has been the subject of a number of studies.[1–6] Irradiation with X-rays or ultraviolet or heating in oxygen or magnesium vapor will produce optical absorption bands in MgO crystals which contain no intentionally added foreign ions. Until recently it was assumed that both the luminescence and the optical absorption were associated with defects in the host lattice other than foreign ions. HAXBY[7] and SOSHEA *et al.*[8] have recently presented evidence that certain of the ultraviolet absorption bands are due to Fe^{+3}. We support this hypothesis. WERTZ *et al.*[9, 10] have evidence for changes in valence of various foreign ions in MgO during X-irradiation and other treatments.

We wish to present evidence that certain luminescent effects in relatively high purity MgO are due to the transfer of electrons (and holes) between the iron and chromium impurities and depend on defects such as ion vacancies only in an indirect manner.

2. EXPERIMENTAL PROCEDURE

Crystals studied

Most of the single crystals used in this investigawere grown in this laboratory by L. J. SCHUPP by the arc melt method. In our highest purity crystals (designated *R*–58 and *R*–72) the chief impurity is calcium (<0·05 per cent)* which probably has little effect because of its chemical similarity to magnesium. The following were spectroscopically detected in amounts less than 0·01 per cent: Si, Al, Fe, Ti, Ni, Zn, Sr, Na, K. Using other methods, *R*–58 crystals were found to contain approximately 5 p.p.m. Sr, 8 p.p.m. Fe, 3 p.p.m. Li and 5 p.p.m. Cr.

* The concentrations of impurities are given either as wt. % or p.p.m. by weight.

VOL. 13 1960

Electronic paramagnetic resonance (EPR) *absorption measurements*

Mn^{+2}, V^{+2}, Fe^{+3} and Cr^{+3} ions in MgO give readily observable peaks in the EPR absorption spectrum. We have observed the changes in concentration of these ions in a number of crystals subjected to various treatments. Quantitative data were obtained for the concentrations of Cr^{+3} and Fe^{+3} in some crystals subjected to certain treatments.

The EPR spectra were taken with a Varian spectrometer and 6-in magnet. The crystals were cemented to a quartz rod and oriented so that H_1 was parallel to the [100] axis and H_0 was parallel to the [010] axis. A $CuSO_4 \cdot 5H_2O$ sample was placed in the cavity and the Cu^{+2} absorption line used as a relative intensity standard.

Only one line of each of the complex spectra of Cr^{+3} and Fe^{+3} was used for the quantitative intensity measurements. For Cr^{+3} the "cubic" line was used and for Fe^{+3} the central line of the five line spectrum was used.[10]

Optical absorption measurements

A Beckman model DU spectrophotometer was used for most of the measurements. After excitation with ultraviolet it was sometimes necessary to check to make sure that long-lived phosphorescence did not interfere with the measurement. The measurements in the vacuum ultraviolet were made by Dr. P. D. JOHNSON at the General Electric Research Laboratory, Schenectady, New York.

Detection of luminescence

Most of the measurements were of a qualitative type in which the color of the thermoluminescence was noted when the crystal was placed on a hot plate at about 300–400°C after being irradiated with light from a 4-W germicidal lamp.* A small quartz spectrograph was used to measure the emission in a few cases.

Glow curves

Thermoluminescent glow curves were obtained for a few crystals. The crystals were heated at a uniform rate and the light output measured with a 1P28 photomultiplier.

* This low pressure mercury discharge lamp produces principally 2537 Å (4·9 eV) radiation.

3. EXPERIMENTAL RESULTS AND INTERPRETATION

Preliminary results of EPR measurements on a number of samples suggested that neither the manganese nor vanadium were directly involved in the observed thermoluminescence. The amount of Mn^{+2} remained remarkably constant under quite a variety of treatments. No correlation seemed to exist either between the amount of V^{+2} present and the luminescence or between the valence change of the vanadium and the luminescence.

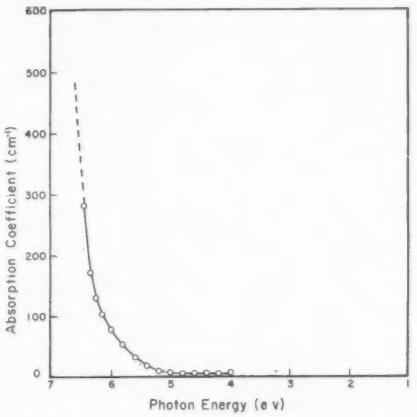


FIG. 1. The absorption coefficient vs. photon energy of an MgO:3%Fe_2O_3 crystal which was fired 6 hr at 1100°C in hydrogen in order to reduce the iron to the divalent state. The sharp rise in absorption at about 6·7 eV we attribute to the Fe^{+2}.

Depending on the previous treatment the amount of Fe^{+3} could either be increased or decreased during ultraviolet irradiation and showed the reverse change during thermoluminescence. The same was observed with Cr^{+3}. These preliminary results suggested the luminescence might be correlated with these changes in valence.

The positions of the energy levels of Fe^{+3}, Fe^{+2}, Cr^{+3} and Cr^{+2} in the MgO lattice were determined by optical absorption. HAXBY[7] showed the correlation between the amount of Fe^{+3} in the lattice and the 4·3 and 5·7 eV absorption peaks.

We found that by hydrogen firing an MgO:3%Fe_2O_3† crystal for six hr at 1100°C all of

†This notation means, "grown from a melt to which 3 per cent by weight Fe_2O_3 was added". The amount of foreign ion in the crystals usually is about one-tenth of this amount.

the 4·3 and 5·7 eV absorption could be removed. Fig. 1 shows the absorption edge at approximately 6·7 eV which we attribute to Fe^{+2}. That the iron was not removed by the hydrogen firing but merely reduced was shown by observing the return of the 4·3 and 5·7 eV absorption after heating the crystal briefly in air.

Fig. 2 shows the absorption of an MgO: 0·1%Cr_2O_3 crystal after hydrogen firing 5 hr at

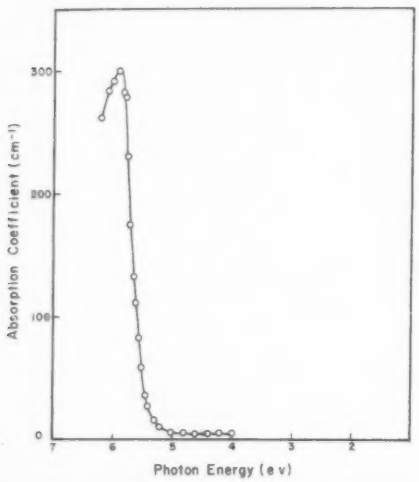


FIG. 2. The absorption coefficient vs. photon energy of an MgO:0·1%Cr_2O_3 crystal which was fired 5 hr at 1200°C in hydrogen to reduce the iron present. This treatment does not reduce the chromium. The 5·9 eV peak is due to Cr^{+3}.

1200°C. A strong peak at 5·9 eV is observed. This crystal contains some iron although none was intentionally added. Our EPR results show that the iron is reduced to Fe^{+2} but that the amount of Cr^{+3} present is a maximum after hydrogen firing. The absence of any 4·3 eV absorption confirms that no Fe^{+3} is present. The occurrence of this absorption at 5·9 eV which we attribute to Cr^{+3} is close to the 5·7 eV Fe^{+3} absorption maximum. It may explain the shifts in the peak observed in this region which were reported by SOSHEA *et al.*[8]

The chromium can be reduced to Cr^{+2} only under unusually strong reducing conditions. EPR measurements show that even firing in magnesium vapor does not reduce an appreciable fraction although such firing will reduce the iron present in the crystal to the metallic state. The presence of metallic iron is inferred from valence studies carried out by R. P. TAYLOR of this laboratory using wet chemical methods. It is also suggested by the fact that crystals containing more than 100 p.p.m. of iron are attracted by a magnet after firing in Mg vapor. We have been able to produce Cr^{+2} in high purity MgO crystals (total chromium about 3 p.p.m.) by passing direct current through the crystals at high temperature. Fig. 3 shows the

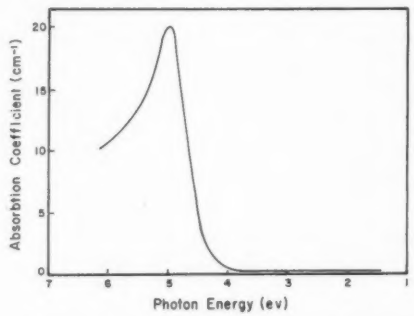


FIG. 3. The absorption coefficient vs. photon energy for a high purity MgO crystal which has been treated by the passage of direct current at high temperature for a few minutes. We attribute the band at 5 eV to Cr^{+2}.

absorption spectrum after this treatment. The band at 5 eV which we think is due to Cr^{+2} has not been observed previously by optical absorption. PERIA[11, 12] reports a peak in the photoconductivity spectrum at 5·05 eV. He states that the charge carriers arising from this center are electrons. This is consistent with the change in valence from Cr^{+2} to Cr^{+3}.

The positions of the absorption bands associated with these ions are tabulated in Table 1.

Table 1

Ion	Optical absorption
Fe^{+2}	6·7 eV
Fe^{+3}	4·3 & 5·7 eV
Cr^{+2}	5 eV
Cr^{+3}	5·9 eV

In order to explain the color of the luminescence which is associated with transitions at these four

ions we propose the band model shown in Fig. 4. The large Stokes shift between the occupied (Fe^{+2} and Cr^{+2}) and the unoccupied (Fe^{+3} and Cr^{+3}) levels is the result of the large polarizability of MgO. The excited states of the centers are sufficiently close to the valence and conduction bands that the carriers can be thermally excited into the bands even at room temperature. If we accept the value for the band gap of 8·7 eV given by REILING and HENSLEY[13] the energy of the photons produced during luminescence is roughly the difference between 8·7 eV and the optical absorption energy. The solid arrows in Fig. 4 indicate the optical absorption energies. The dashed arrows show the electron and hole transitions giving rise to luminescence.

With our proposed model in mind we sought out crystals in which one of four reactions dominated the others. The processes which produce the luminescence, and the energy and color of the

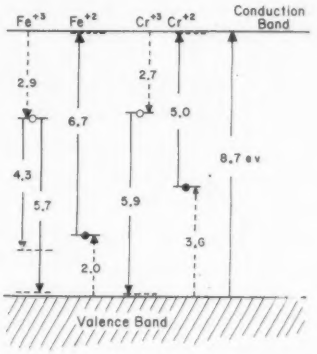


FIG. 4. An energy level scheme showing the energy levels introduced in magnesium oxide by di- and trivalent iron and chromium. The solid arrows show the transitions causing optical absorption. The dashed arrows show the transitions which produce luminescence. The trapping states for electrons and holes which can be emptied by heating the crystal are not shown.

Table 2. The concentration of Fe^{+3} *and* Cr^{+3} *in four* MgO *crystals after various treatments. The arrow shows which change in concentration produced the luminescence indicated in the reaction.*

Sample no.	Reaction	Treatment	Concentration of Fe^{+3}	Concentration of Cr^{+3}
1.	Fe^{+3}+free electron→ Fe^{+2}+2·9 eV (blue)	Ultraviolet	78×10^{15} cm^{-3}	12×10^{15} cm^{-3}
		Partial thermoluminescence	106 ↓	5
		Complete thermoluminescence	48	6
2.	Fe^{+2}+free hole→ Fe^{+3}+2·0 eV (red)	Ultraviolet	290 ↓	199
		1st Thermoluminescence peak	440 ↓	194
		2nd Thermoluminescence peak	570 ↓	206
3.	Cr^{+3}+free electron→ Cr^{+2}+2·7 eV (blue)	Ultraviolet	78	85 ↓
		1st Thermoluminescence peak	109	74 ↓
		2nd Thermoluminescence peak	116	69 ↓
		Complete thermoluminescence	96	47
4.	Cr^{+2}+free hole→ Cr^{+3}+3·6 eV (blue)	Ultraviolet	65	256 ↓
		1st Thermoluminescence peak	96	269 ↓
		2nd Thermoluminescence peak	92	228 ↓
		3rd Thermoluminescence peak	129	276

light are shown in Table 2. The concentrations given in the last column of Table 2 are the EPR intensities multiplied by a calibration factor. The calibration factor was obtained by comparing the maximum Fe^{+3} or Cr^{+3} EPR intensity from a given crystal (highly oxidized) with the total amount of Fe or Cr in that crystal as determined by chemical means. Therefore, the numbers in the table are really the upper limits on the concentrations instead of the concentrations. The relative amounts of Fe^{+3} (or of Cr^{+3}), however, are in error by no more than 10 per cent (due to "noise" in the EPR spectrum).

The first sample mentioned in Table 2 was of high purity (from melt *R*–72) and had been treated by passing d.c. through it at a high temperature resulting in a reduction of most of the iron and chromium. During ultraviolet excitation evidently both holes and electrons are freed from iron and chromium and become trapped elsewhere. The glow curves for this crystal are shown in Fig. 5.

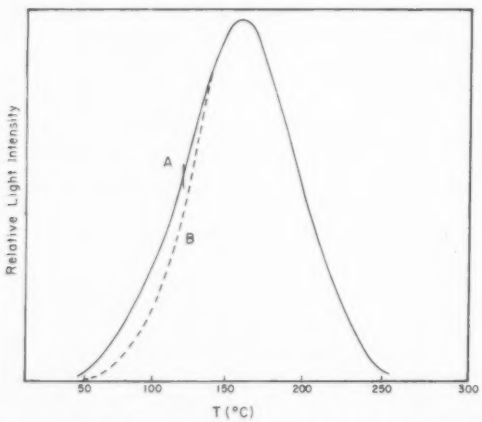


FIG. 5. The light as measured by a 1*P*28 photomultiplier as a function of crystal temperature for a uniform heating rate of 0·23°/sec. The (*R*–72) crystal had been treated by the passage of direct current at high temperature. It was irradiated with 2537 Å light before being heated to produce the glow curve *A*. Curve *B* was obtained by heating the crystal only to 115° after the irradiation, cooling to room temperature and then heating while recording the light output.

The complete curve (*A*) with a strong peak at 160°C and a very weak one at 90°C was made by continuous heating at a rate of 0·23°/sec. The crystal was then irradiated again and heated to 115°C, cooled to room temperature and the EPR spectrum measured. In Table 2 the resulting changes are recorded after "Partial Thermoluminescence". Without further irradiation it was then heated to 255°C (curve *B*), cooled to room temperature and the EPR spectrum measured again. The results are tabulated after "complete thermoluminescence" in Table 2. The Cr^{+3} decreases during the first part of the thermoluminescence while the Fe^{+3} increases. This implies that the Cr^{+3} must have a larger capture cross section for electrons and the Fe^{+2} for holes. During the complete thermoluminescence the Fe^{+3} picks up most of the electrons liberated at the higher temperature. Evidently the majority of the holes are liberated during the first step. The 90°C peak may be associated with the untrapping of holes. The luminescence would be red corresponding to hole capture by Fe^{+2}. No red luminescence was visible. The relatively intense blue luminescence probably masks it. During the second step the emission consists of a blue band between 2·7 and 3·4 eV, in good agreement with the predicted 2·9 eV. It should be pointed out that the position (temperature) of the glow peaks is only dependent on the nature of the hole and electron traps and the rate of heating. The position is not dependent on the nature of the luminescent center.

The second sample, also of high purity but from

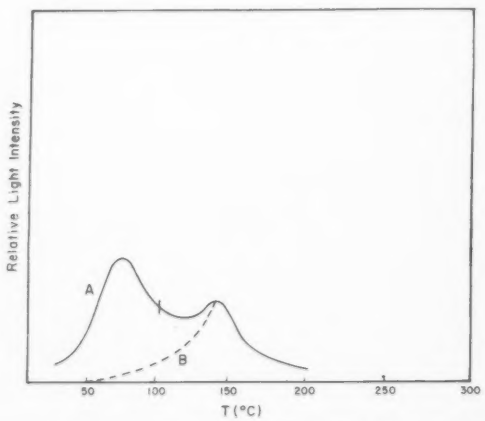


FIG. 6. The glow curve of a high purity (*R*–58) untreated MgO crystal. Curve *A* was made during continuous heating following 2537 Å irradiation. Curve *B* was obtained after a preliminary heating to 110°C to remove the carriers causing the lower temperature peak.

VOL. 13 1960

a different melt (*R*–58), was not treated other than to cleave it from a larger piece. The concentrations of both Fe^{+3} and Cr^{+3} are much higher. Fig. 6 shows the glow curve made in one continuous heating (curve *A*) and that made to determine the valence changes which occurred during the second peak (curve *B*). As shown in Table 2, only holes freed from the Fe^{+3} during ultraviolet excitation and returned during the thermoluminescence appear to be involved. The two peaks in the glow curve suggest the presence of two different types of hole traps. The luminescence consists only of red light in a rather narrow band at 2·0 eV. This is consistent with our model.

The third sample was from the same melt as the second (*R*–58) but was treated by passage of direct current at high temperature. Fig. 7 shows the glow curve obtained with continuous heating (curve *A*) and that made to check on the changes occurring during the highest temperature peak (curve *B*). The shoulder on the first peak is evidence that we are not separating all the effects. With the glow curve apparatus it was not possible to obtain temperatures high enough to untrap all the carriers. The last entry in Table 2 under sample 3 gives the concentrations of Fe^{+3} and Cr^{+3} after heating the crystal on a hot plate at 300–400°C for 15 min.

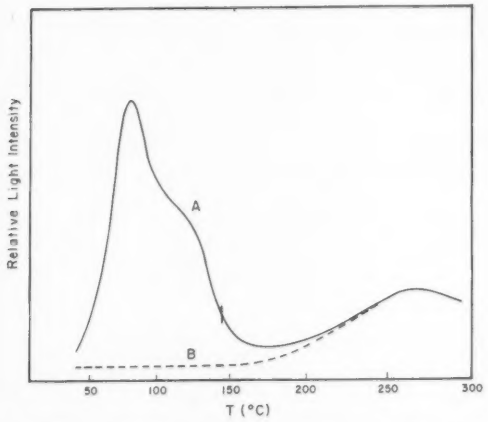


FIG. 7. The glow curve of a high purity (*R*–58) MgO crystal which had been treated by passage of direct current at high temperature. Curve *A* was obtained by continuous heating following 2537 Å irradiation. Curve *B* was obtained after a preliminary heating to remove the carriers causing the lower temperature peak.

Only blue luminescence was observed. The slight amount of red which our model predicts we might see during the first peak was probably masked by the relatively intense blue luminescence. The Cr^{+3} concentration dropped during each step. The light emitted during the first peak consists of a broad band from 2·75 to 4·3 eV. The light produced during the highest temperature peak is weaker and extends approximately from 2·75 to 3·45 eV. The fact that the Fe^{+3} concentration increased during the first peak is evidence for the presence of free holes. If some of these holes are trapped by the Cr^{+2}, we should expect some 3·6 eV photons. The emission during the first peak extending to 4·3 eV might be explained in this way. During the highest temperature peak and at higher temperatures (the spectrum was taken by heating to about 400°C following the heating and cooling carried out to obtain the spectrum of the first peak) there is little evidence for the presence of holes. The observed spectrum extending from 2·75 to 3·45 eV is thus in fair agreement with the 2·7 eV peak predicted by the model for electron capture by Cr^{+2}.

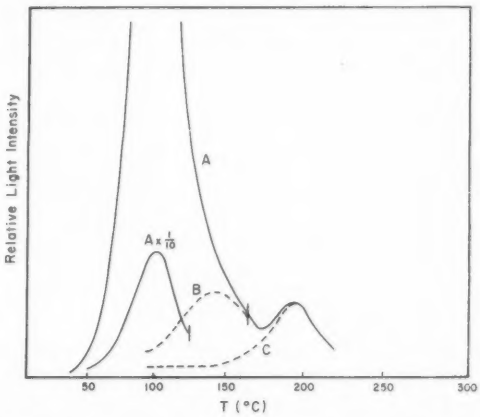


FIG. 8. The glow curves for a moderate purity MgO crystal which was treated by oxygen firing one hour at 1200°C. Three curves taken to separate the three types of traps are shown.

The fourth sample is from a different melt (*P*–5) and contains a relatively larger amount of chromium. It was heated 1 hr in oxygen at about 1200°C. The glow curves in Fig. 8 are shown for continuous heating (curve *A*) and for heating

schedules used to separate the changes occurring during the three peaks (curves *B* and *C*). The first and third peaks indicate that hole capture by Cr^{+2} is occurring while the second peak seems to involve electron capture by Cr^{+3}. The thermoluminescence of sample 4 was too weak to be detected by our spectrograph. During thermoluminescence the crystal first appears purple and then red. The purple color is the result of both red and blue emission occurring simultaneously. The red should arise from hole capture by Fe^{+2} while the blue is the result of hole capture by Cr^{+2}. This is the least convincing of the four experiments, but thus far we have been unable to find any crystal in which the only significant change was hole capture by Cr^{+2}.

The results of the measurements on these four crystals offer support for the model we are proposing. Measurements on many other crystals have failed to produce any effect which is not satisfactorily explained in terms of this model.

We have thus far said nothing about the exact nature of the hole and electron traps. At least two optical absorption bands have been observed (2·3 eV and 3·6 eV) which we may attribute to trapped holes and electrons. Fig. 9 shows the optical absorption spectrum of a moderately high purity (*P*–5) neutron-irradiated sample. Peaks at 1·3, 2·2, 3·0 and 3·5 eV are seen.

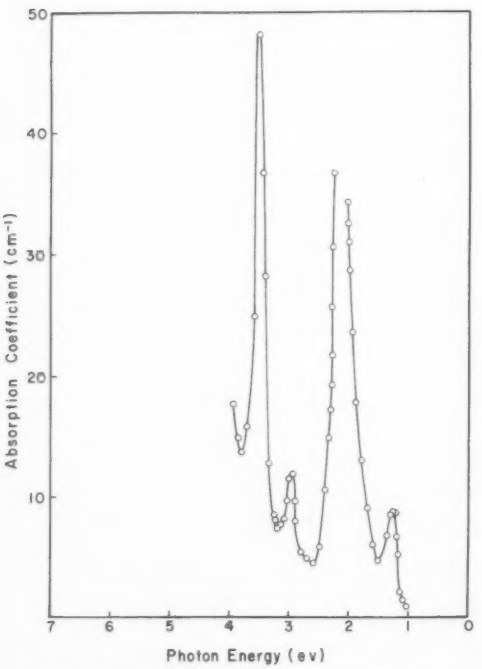


FIG. 9. The absorption coefficient vs. photon energy for a neutron irradiated moderate purity (*P*–5) MgO crystal. The integrated neutron dose was $1{\cdot}2 \pm 0{\cdot}2 \times 10^{20}$ nvt.

An analysis of the glow peaks of the type described by RANDALL and WILKINS[14] in which we assume no retrapping, gives values for the activation energy of 0·2–1 eV and frequency factors of 10^5 to 10^{12} sec^{-1}. Such calculations are probably of little significance, however, for there is no reason to believe that retrapping does not occur. The complexity of the glow curves indicates that at least four types of traps exist. Magnesium vacancies should trap either one or two holes and thus give rise to two types of trapping centers. The observation by ORTON[15] of Fe^{+1} in crystals which have been ultraviolet irradiated makes Fe^{+2} a good possibility for an electron trap. The identification of the traps with the optical absorption bands is one of the attractive problems remaining in the study of magnesium oxide.

4. SUMMARY

1. Optical absorption studies have shown the position of the energy levels due to Fe^{+2}, Fe^{+3}, Cr^{+2} and Cr^{+3}.

2. By assuming the band gap in MgO is 8·7 eV (after REILING[13]) the energy of the luminescence due to recombination of free holes with Fe^{+2} and Cr^{+2} and free electrons with Fe^{+3} and Cr^{+3} is predicted.

3. EPR studies have made it possible to select crystals in which only one of the four processes takes place during some part of the thermoluminescence which is carried on after excitation with 4·9 eV (2537 Å) photons.

4. The spectrum of the light emitted during this one process has been measured for each of the four cases and compared with that predicted by the model. The agreement is good.

5. The thermoluminescent glow curves offer information about the hole and electron traps. No identification of the traps has been made.

Acknowledgements—It is our pleasure to thank our colleagues in the Lamp Research Laboratory for their contributions to this paper. We also wish to express

our appreciation to the groups at the University of Minnesota and the University of Missouri who have contributed so much to the knowledge of magnesium oxide.

REFERENCES

1. Bube H. and Stipp K. F., *J. Chem. Phys.* **20**, 193 (1952).
2. Prener J. S., *J. Chem. Phys.* **21**, 160 (1953).
3. Yamaka E., *Phys. Rev.* **96**, 293 (1954).
4. Saksena B. D. and Pant L. M., *Proc. Phys. Soc., Lond.* **67**, 811 (1954).
5. Saksena B. D. and Pant L. M., *Curr. Sci.* **23**, 393 (1954).
6. Saksena B. D. and Pant L. M., *J. Chem. Phys.* **23**, 989 (1955).
7. Haxby B. V., *A Study of Color Centers in Magnesium Oxide*, Ph.D. Thesis, Univ. of Minn. (1957).
8. Soshea R. W., Dekker A. J. and Sturtz J. P., *J. Phys. Chem. Solids* **5**, 23 (1958).
9. Wertz J. E. and Auzins P., *Phys. Rev.* **106**, 484 (1957).
10. Wertz J. E., Auzins P., Griffiths J. H. E. and Orton J. W., *Disc. Faraday Soc.* **26**, 66 (1958).
11. Peria W. T., *Optical Absorption and Photoconductivity in Magnesium Oxide Crystals*, Ph.D. Thesis, Univ. of British Columbia (1957).
12. Peria W. T., *Phys. Rev.* **112**, 423 (1958).
13. Reiling G. H. and Hensley E. B., *Phys. Rev.* **112**, 1106 (1958).
14. Randall J. T. and Wilkins M. H. F., *Proc. Roy. Soc.* A**184**, 366 (1945).
15. Orton J., private communication.

VOL.
13
1960

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 13. pp. 132–144. Printed in Great Britain.

A HEAT CAPACITY STUDY OF THE MOVEMENT OF THE AMMONIUM ION IN AMMONIUM STANNICHLORIDE AND STANNIBROMIDE BY COMPARISON OF THE HEAT CAPACITIES OF THE AMMONIUM, RUBIDIUM AND POTASSIUM SALTS

R. G. S. MORFEE, L. A. K. STAVELEY, S. T. WALTERS and D. L. WIGLEY

The Inorganic Chemistry Laboratory, Oxford University, Oxford, England.

(*Received* 3 *July* 1959)

Abstract—The heat capacity of the following six salts has been measured from ~20°K to ~300°K: ammonium, rubidium and potassium stannichlorides, ammonium, rubidium and potassium stannibromides. By comparison of the heat capacities of the ammonium and rubidium salts (which are isomorphous), that part of the heat capacity due to the torsional oscillations of the ammonium ions has been separated, and for both the stannichloride and the stannibromide is found to be decreasing as room temperature is approached from below, reaching at 298°K a value roughly halfway between that for free rotation and for classical torsional oscillations. This implies that the ammonium ions are restricted rotators prevented from freely rotating by comparatively low energy barriers. The bearing of this on the charge distribution in the stannihalide ions has been considered.

Gradual transitions have been found in ammonium stannibromide and potassium stannichloride. Whereas the latter is apparently a simple λ-type transition with an entropy change of ~1·1$_3$ e.u., that in the stannibromide has an entropy change of ~4·5 e.u. and appears to be a transition of some complexity. The possibility is considered that the transition in potassium stannichloride might be due to the availability to the anion of two alternative orientations.

ALTHOUGH the ammonium ion in simple ammonium salts does not undergo free three-dimensional rotation at room temperature, an ammonium ion could undergo such motion, or a near approach to it, when present in a lattice in which the field of force acting on the ion is sufficiently uniform to make the potential barrier to rotation small. Such a field of force might be expected in salts of high lattice symmetry where the anions are polyatomic and not too highly charged. In this paper we describe experiments to investigate the freedom of the rotational movement of the ammonium ions in the lattices of ammonium stannichloride and ammonium stannibromide. These salts were chosen since at ordinary temperatures they have anti-fluorite structures in which each ammonium ion, being surrounded tetrahedrally by four stannihalide ions, has as its nearest neighbours twelve equidistant halogen atoms. These atoms are very symmetrically disposed about the cation, since they lie on or parallel to the crystal axes and since the tin-halogen bond length is very nearly one quarter of the length of the side of the unit cell. Moreover, as the total charge on the stannihalide ion is two, the effective charge on any one halogen cannot be large. These factors should favour a comparatively small barrier to the rotation of the ammonium ion.

The principle of our method is simply to compare the molar heat capacities C_p of the corresponding ammonium and rubidium salts, which are isomorphous and have almost the same lattice dimensions. The actual values for the sides of the cubic unit cell are $(NH_4)_2SnCl_6$, 10·038 Å; Rb_2SnCl_6, 10·099 Å; $(NH_4)_2SnBr_6$, 10·59 Å; Rb_2SnBr_6, 10·64 Å.[1] It is reasonable to suppose that the following contributions to C_p will therefore not differ significantly for the ammonium and

VOL. 13 1960

rubidium salts: (i) $C_p - C_v$; (ii) the contribution from the internal vibrations of the anion; and (iii) that from the torsional oscillations of the anion. Consequently, at temperatures sufficiently high for the lattice vibrations to contribute the maximum limiting amount to C_p (a state which should virtually be reached well below room temperature), the quantity ΔC_p, $= C_p$ (NH_4 salt) $- C_p$ (Rb salt) $- 2C_{int}$ (NH_4), should be the heat capacity contribution from the torsional or rotational movement of the ammonium ion. C_{int} (NH_4) is the relatively small contribution from the internal vibrations of the ammonium ions, which can be calculated from the frequency assignment of WAGNER and HORNIG[6].

Since the lattice dimensions of the ammonium and rubidium salts do in fact differ, if only slightly, we have also made measurements on the corresponding potassium salts. The length of the side of the cubic unit cell of K_2SnCl_6 is 9·983 Å. It is possible that K_2SnBr_6, almost alone among salts of this type, is not isomorphous with the rest, for while KETELAAR *et al.*[2] considered it to be cubic, MARKSTEIN and NOWOTNY[3] concluded that it has a slightly modified, tetragonal structure.

A similar investigation to this on the ammonium, rubidium and potassium salts of tetraphenylboron has already been described.[4]

EXPERIMENTAL

The heat capacities of the six salts were measured from ~20°K to ~300°K, using the calorimeter described by DAVIES and STAVELEY[4]. We have drawn attention elsewhere[5] to an unusual phenomenon encountered in these measurements, namely that the heat capacities of these salts at higher temperatures is reduced by cooling, and only attains reproducible values after repeated cooling. The C_p figures recorded in this paper are throughout the final reproducible values.

A good sample of $(NH_4)_2SnCl_6$ was already available. K_2SnCl_6 was prepared by mixing concentrated solutions of KCl and $SnCl_4$ in stoichiometric proportions in the cold and slowly evaporating the resulting solution in a vacuum desiccator. Rb_2SnCl_6 (which has a low solubility) was obtained by adding a hot solution of $SnCl_4$ in hydrochloric acid to a hot acid solution of RbCl. The rubidium stannichloride separated on cooling, and further crystallisation was allowed to proceed in a vacuum desicator. $(NH_4)_2SnBr_6$ was readily obtained as a yellow precipitate by adding a solution of ammonium bromide in water to the stoichiometric quantity of molten stannic bromide. Recrystallisation from water at room temperature gave large crystals. The preparation of stable samples of the potassium and rubidium salts gave a little more trouble. It was found that a solution made from the stoichiometric proportions of KBr and $SnBr_4$ deposited potassium bromide on evaporation. Large crystals of K_2SnBr_6 were however obtained from a solution containing equimolar amounts of KBr and $SnBr_4$. The rubidium salt was prepared by adding rubidium bromide in stoichiometric proportion to a solution of stannic bromide in hydrobromic acid.

All salts were separated from the mother liquor by centrifuging, dried *in vacuo* over P_2O_5 and finally lightly crushed and sieved to give samples for the calorimeter of approximately the same particle size.

All salts were analysed for the halogen content by first hydrolysing them to precipitate the tin as SnO_2, and then determining the halogen content of the filtrate gravimetrically as the silver halide. The tin content was determined by one of the following methods: (i) iodimetrically, after its reduction to the stannous state by antimony in an inert atmosphere (K_2SnCl_6, Rb_2SnCl_6, $(NH_4)_2SnCl_6$, $(NH_2)_2SnBr_6$), (ii) gravimetrically as SnO_2 (K_2SnBr_6), (iii) by the electrolytic method of SCHOCH and BROWN[7], modified by ENGELENBURG[8], (Rb_2SnBr_6). The ammonium content of the two ammonium salts was found by distilling off the ammonia with alkali into excess standard acid. The analytical results are shown in Table 1. It will be noticed that there is a tendency for the halogen content to be below the theoretical value. Mr. G. SAVILLE has more recently made careful analyses in this laboratory of K_2SnCl_6 and $(NH_4)_2SnCl_6$ prepared under various

Table 1. *Analytical results for the salts studied calorimetrically*

		$(NH_4)_2SnCl_6$	K_2SnCl_6	Rb_2SnCl_6	$(NH_4)_2SnBr_6$	K_2SnBr_6	Rb_2SnBr_6
% NH_4	theory	9·82	—	—	5·69	—	—
	expt.	9·81	—	—	5·70	—	—
% Sn	theory	32·30	28·98	23·63	18·70	17·55	15·44
	expt.	32·25	28·8	23·55	18·73	17·55	15·55
% Halogen	theory	57·88	51·92	42·34	75·59	70·89	62·35
	expt.	57·60	51·76	41·91	75·52	71·11	62·29

conditions and finds that these two salts are always slightly deficient in chlorine.

RESULTS

The heat capacity results are recorded in Table 2, and smoothed values at regular temperature intervals are given in Table 3. Plots of C_p against temperature for Rb_2SnCl_6 and K_2SnBr_6 show no signs of any anomaly. $(NH_4)_2\,SnCl_6$ shows a slight anomaly between 235 and 245°K (see Fig. 1). The eutectic temperature for the system of this salt with water is not known. The anomaly might be due to

Table 2. Measured values of C_p, the molar heat capacity in cal/mole deg.

1 calorie = 4·184 absolute joules. T = temperature in °K (0°C = 273·16°K).

Ammonium stannichloride

T	C_p	T	C_p	T	C_p
21·49	6·18	111·40	47·47	220·66	60·18
23·55	7·37	116·79	48·81	223·00	60·32
26·10	8·78	122·31	49·87	229·43	60·70
30·04	11·36	127·72	50·95	234·90	60·95
35·02	14·71	133·02	52·07	235·38	61·01
40·78	18·44	138·05	52·92	240·68	61·00
48·16	22·97	143·27	53·76	245·63	61·07
55·48	27·35	149·05	54·54	246·46	61·14
60·82	30·34	154·86	55·50	252·88	61·43
65·02	32·17	158·74	55·97	254·31	61·41
68·79	33·77	162·59	56·39	258·72	61·58
73·56	35·53	169·50	56·86	263·06	61·72
78·79	37·53	175·65	57·40	271·68	62·13
83·79	39·44	181·74	58·22	280·66	62·50
88·60	41·08	187·63	58·56	289·72	63·07
93·41	42·44	193·79	58·89	298·11	63·15
97·81	43·97	194·58	58·90		
101·60	44·92	202·84	59·31		
106·16	45·98	211·85	59·81		

Potassium stannichloride

T	C_p	T	C_p	T	C_p
21·54	7·42	99·55	38·00	185·17	49·84
23·47	8·44	100·94	38·35	191·66	50·53
25·77	9·81	104·97	39·15	198·29	50·94
29·39	11·57	106·77	39·52	205·07	51·40
35·30	14·87	110·78	40·31	211·94	51·94
41·89	18·19	118·08	41·47	218·83	52·54
48·58	21·51	124·93	42·79	225·76	53·11
55·43	24·51	130·70	43·64	234·23	54·16
62·00	27·19	136·57	44·40	241·53	54·90
67·06	28·91	140·85	45·23	249·78	55·54
71·14	30·39	146·78	46·01	256·63	57·51
76·10	31·97	153·79	46·58	259·74	63·70
81·47	33·74	160·24	47·37	263·88	62·25
86·00	35·06	166·41	48·11	270·21	52·29
89·76	35·74	172·63	48·74	271·22	52·24
93·74	36·59	178·94	49·53	280·03	52·25
				288·83	52·44
				297·58	53·01

Table 2. Measured values of C_p, the molar heat capacity in cal/mole deg (continued).

1 calorie = 4·184 absolute joules. T = temperature in °K (0°C = 273·16°K).

Rubidium stannichloride

T	C_p	T	C_p	T	C_p
21·90	8·82	98·78	38·33	190·79	49·20
23·85	10·24	104·81	39·35	197·70	49·54
26·27	12·34	110·95	40·30	204·75	49·89
29·78	14·15	116·83	41·58	211·70	50·44
35·15	17·50	122·22	42·24	218·79	50·80
41·40	20·36	128·06	43·22	226·05	51·21
48·21	23·44	134·16	44·02	233·50	51·21
55·26	26·26	140·08	44·67	241·15	51·71
61·43	28·54	145·89	45·37	248·90	52·06
66·79	30·41	152·21	46·02	257·17	52·23
71·65	31·82	158·62	46·56	266·15	52·99
76·76	33·14	164·92	47·19	275·88	53·13
82·17	34·56	171·24	47·72	286·09	53·60
87·79	36·20	177·52	48·22	296·53	54·01
93·18	37·25	183·85	48·69		

Ammonium stannibromide

T	C_p	T	C_p	T	C_p
21·37	9·87	109·40	62·69	149·97	61·92
21·56	9·81	110·36	62·78	151·02	61·47
23·18	10·59	113·11	63·77	152·63	62·35
24·01	11·06	114·37	63·79	153·13	62·21
25·45	12·02	116·91	64·81	155·21	62·25
26·23	12·43	118·27	64·72	156·85	62·85
28·64	14·29	120·20	65·43	157·27	62·50
29·28	14·58	123·49	65·75	160·09	62·96
32·71	17·06	124·60	66·24	163·07	62·71
33·84	17·82	127·99	66·27	164·10	61·62
37·00	19·97	128·91	67·17	166·25	61·60
39·81	21·87	131·69	67·13	167·46	61·35
46·39	26·39	132·68	66·85	168·40	61·58
53·40	31·17	132·95	67·72	169·83	61·65
59·90	35·90	134·48	67·50	173·08	61·29
64·66	39·11	135·78	67·36	180·84	61·33
68·60	41·39	137·01	68·68	185·82	61·55
72·63	43·98	137·67	68·32	192·24	61·41
76·84	46·60	138·23	68·43	198·56	61·61
81·80	50·12	139·45	68·05	204·76	61·76
87·15	52·63	140·65	69·05	211·13	61·91
95·35	57·60	141·84	69·23	217·48	62·36
98·72	58·83	143·01	70·93	223·94	62·47
102·01	60·05	143·65	88·86	229·07	62·40
105·63	61·33	147·27	61·32	236·82	62·36
106·22	61·58	148·75	61·83	245·74	62·75
				252·09	62·71
				260·33	63·01
				268·69	63·24
				286·84	63·56
				295·98	63·89

VOL. 13 1960

Table 2. Measured values of C_p, the molar heat capacity in cal/mole deg (continued).

1 calorie = 4·184 absolute joules. T = temperature in °K (0°C = 273·16°K).

Potassium stannibromide

T	C_p	T	C_p	T	C_p
21·32	10·80	102·50	45·91	194·47	55·09
23·22	12·20	107·60	46·64	202·37	55·61
25·64	13·88	113·14	47·47	210·09	55·83
29·39	16·62	119·17	48·43	217·75	56·28
36·73	21·92	125·00	49·22	225·37	56·64
44·99	26·99	130·72	49·78	232·95	57·07
53·11	31·47	136·54	50·53	243·79	56·76
60·79	34·90	144·72	51·58	253·54	57·49
64·71	36·42	151·65	52·17	261·96	58·03
70·23	38·23	158·47	52·30	269·78	57·63
76·30	39·96	165·22	52·85	277·98	58·28
81·96	41·55	171·85	53·32	286·30	58·46
88·52	43·02	178·83	53·79	292·97	59·11
97·21	45·04	186·50	54·20	301·34	58·74

Rubidium stannibromide

T	C_p	T	C_p	T	C_p
21·10	12·34	90·35	44·20	194·16	51·73
22·50	14·93	92·40	44·39	201·11	51·76
24·55	16·78	93·24	44·30	208·20	51·91
27·88	19·66	94·31	44·43	214·85	52·17
33·33	23·49	96·20	44·77	220·97	52·42
39·95	27·54	97·98	44·67	227·22	52·59
46·84	30·48	98·06	44·95	235·80	52·87
54·87	34·43	100·80	45·20	236·23	52·86
62·56	37·47	103·17	45·38	244·04	53·10
68·71	39·19	108·63	46·00	251·65	53·34
73·34	40·50	114·18	46·74	252·34	53·10
73·88	40·72	119·84	47·19	259·33	53·54
77·39	41·51	127·75	47·93	261·80	53·42
78·33	42·04	135·52	48·49	267·29	53·44
81·62	42·71	143·63	49·16	270·32	53·54
81·90	42·64	150·65	49·80	275·74	54·01
84·34	43·43	157·78	50·14	278·50	53·89
85·67	43·56	164·97	50·44	284·60	53·93
86·38	43·99	172·19	50·83	286·14	54·22
88·38	44·00	180·05	50·96	293·42	54·33
89·59	44·03	186·75	51·30	293·69	54·42
				302·16	54·65

crystallisation at the eutectic point of solution occluded within the crystals, although equilibrium appeared to be established as rapidly in the anomalous region as elsewhere. However, whatever the cause of the anomaly, the associated entropy effect is very small, and it is not surprising that FREYMANN[9] from a study of the dielectric constant of this salt down to liquid air temperatures reported no transition. By contrast, K_2SnCl_6 has a λ -point (Fig. 2), C_p reaching its maximum at $262{\cdot}1 \pm 0{\cdot}05$°K. The determination of the entropy change associated with this transition was assisted by using the results for $(NH_4)_2SnCl_6$ for comparison. It is usually not easy from a single

Table 3. Smoothed values of the molar heat capacity C_p, in cal/mole deg, *at regular intervals of the absolute temperature T.*

1 calorie = 4·184 absolute joules. 0°C = 273·16°K.

T	$(NH_4)_2SnCl_6$	K_2SnCl_6	Rb_2SnCl_6	$(NH_4)_2SnBr_6$	K_2SnBr_6	Rb_2SnBr_6
20	5·30	6·70	7·25	9·27	9·80	10·07
25	8·16	9·34	11·27	11·66	13·50	17·24
30	11·34	12·07	14·29	15·12	17·10	21·26
35	14·68	14·82	17·11	18·64	20·55	24·56
40	17·95	17·48	19·68	22·07	23·85	27·59
45	21·10	19·90	22·04	25·48	26·98	30·24
50	24·14	22·14	24·20	28·91	29·74	32·57
55	27·08	24·28	26·18	32·31	32·28	34·66
60	29·86	26·40	28·00	35·74	34·49	36·50
65	32·23	28·30	29·68	39·10	36·43	38·12
70	34·24	30·02	31·21	42·38	38·17	39·55
75	36·24	31·65	32·71	45·64	39·69	40·94
80	38·12	33·17	34·08	48·84	41·01	42·28
85	39·88	34·56	35·36	51·81	42·28	43·58
90	41·52	35·84	36·52	54·55	43·41	44·06
95	43·06	37·01	38·55	57·07	44·45	44·55
100	44·48	38·11	39·46	59·31	45·40	45·11
110	47·10	40·16	40·31	62·67	46·95	46·21
120	49·45	41·94	41·95	65·32	48·36	47·19
130	51·46	43·54	43·47	66·59	49·64	48·09
140	53·22	44·97	44·71	68·74	50·80	48·89
150	54·76	46·27	45·80	61·65	51·81	49·61
160	56·09	47·40	46·76	62·98	52·68	50·18
170	57·17	48·46	47·63	61·32	53·42	50·65
180	58·08	49·41	48·42	61·32	54·06	51·04
190	58·67	50·30	49·13	61·44	54·70	51·37
200	59·20	51·06	49·77	61·66	55·35	51·69
210	59·69	51·78	50·30	61·93	55·88	52·02
220	60·20	52·62	50·80	62·13	56·31	52·34
230	60·70	53·52	51·24	62·33	56·73	52·64
240	61·04	54·57	51·68	62·53	57·09	52·91
250	61·24	55·75	52·11	62·73	57·44	53·18
260	61·65	64·4	52·54	62·94	57·75	53·45
270	62·06	52·28	52·97	63·16	58·05	53·72
280	62·48	52·24	53·40	63·41	58·34	53·99
290	62·89	52·52	53·84	63·68	58·60	54·27
300	63·30	52·85	54·27	63·97	58·85	54·54

heat capacity curve for a substance as complex as K_2SnCl_6 to decide at what temperature a C_p anomaly may be said to begin and what is the "normal" heat capacity in the transition region. But from Fig. 3, in which $\Delta C_p(NH_4-K) = C_p(NH_4\text{ salt}) - C_p(K\text{ salt}) - 2C_{int}(NH_4)$, is plotted against temperature it appears that the C_p anomaly in the potassium salt begins at about 160°K, and by drawing the dotted curve of Fig. 3 reasonably accurate estimates of the "normal" heat capacity of K_2SnCl_6 can be made over the considerable temperature range covered by the transition. Since this range was about 100°, the entropy of transition was estimated in two parts. Two measurements were made of the energy intake needed to take the salt through the final stages of the transition (from ~248 to ~270°K. By subtracting the "normal" heat intake for this interval the heat absorption associated with the transition (a few per cent only of the whole intake) was

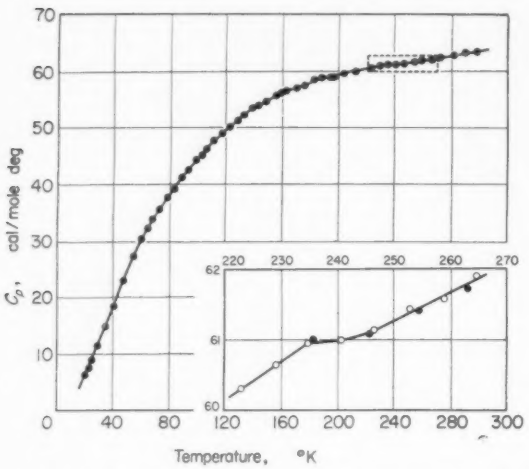


FIG. 1. C_p vs. T for ammonium stannichloride. The inset shows two series of measurements in the region 220–270°K.

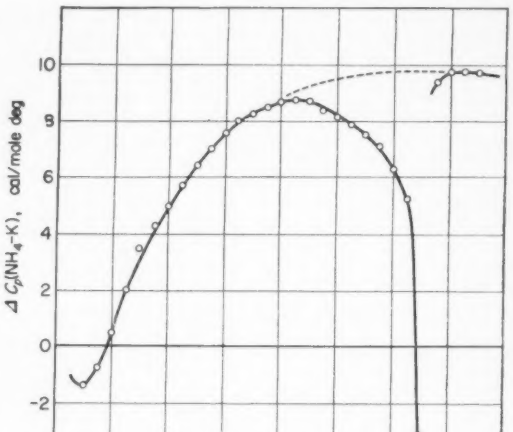


FIG. 3. $\Delta C_p(NH_4-K)$ for the stannichlorides vs. T. $\Delta C_p(NH_4-K) = C_p(NH_4 \text{ salt}) - C_p(K \text{ salt}) - 2C_{int}(NH_4)$.

obtained, and this was assumed for the entropy estimation to be wholly absorbed at 262·1°K. The rest of the entropy increase associated with the interval ~160–248°K was obtained by graphical integration of $\int C_p' d \ln T$, where C_p' is the measured molar heat capacity of K_2SnCl_6 less the "normal" heat capacity, estimated as indicated above. The final result for the entropy of transition was 1·13 cal/mole deg. The error in this may be one or two tenths of a unit, and the figure is therefore not inconsistent with $R \ln 2$.

$(NH_4)_2SnBr_6$ was also found to have a λ-point, with C_p reaching a maximum at 144·8°K. Slightly above the λ-point there is a small but reproducible hump in the heat capacity curve (Fig. 4). In Fig. 5, $C_p(NH_4-Rb)$, $= C_p(NH_4 \text{ salt}) - C_p(Rb \text{ salt}) - 2C_{int}(NH_4)$, is plotted against temperature. The

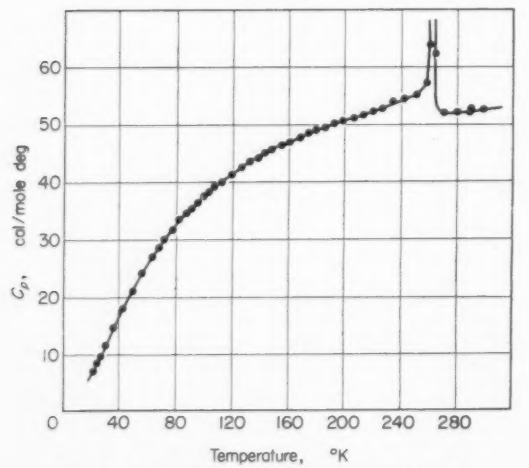


FIG. 2. C_p vs. T for potassium stannichloride.

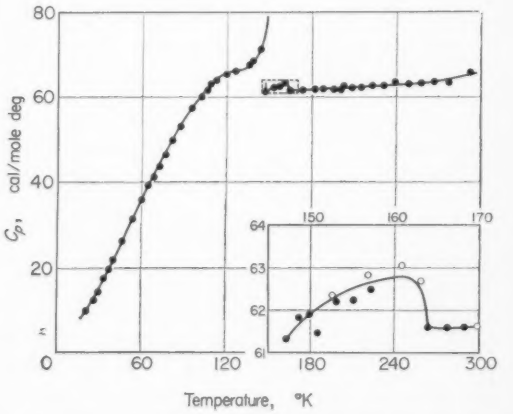


FIG. 4. C_p vs. T for ammonium stannibromide. The inset shows the region 145–170°K (two series of measurements).

heat capacity anomaly extends over a considerable range, and there is some latitude in drawing the dotted curve in Fig. 5 to represent the course of the normal heat capacity of the ammonium salt.

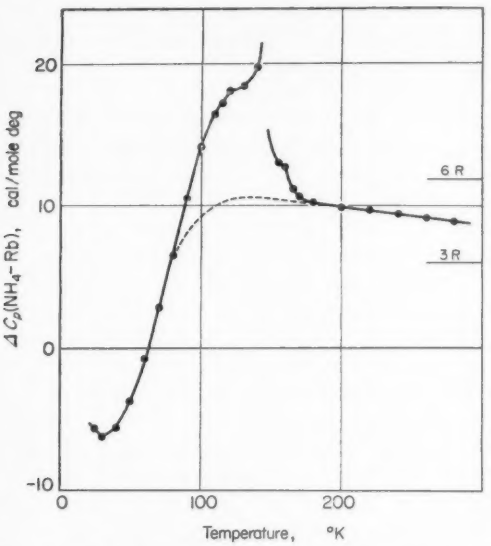


FIG. 5. $\Delta C_p(NH_4-Rb)$ for the stannibromides vs. T. $\Delta C_p(NH_4 - Rb) = C_p(NH_4\,salt) - C_p(Rb\,salt) - 2C_{int}(NH_4)$.

To determine the extra heat intake in the region of maximum C_p, four experiments were carried out in which the salt was heated from a temperature between 120 and 131°K to a final temperature of about 150°K. From the total measured heat intake was subtracted the heat which would have been absorbed, had the heat capacity of the salt up to 144·8°K (the λ-point) followed the extrapolation of the line drawn through the experimental C_p points from 120 to 130°K and, thereafter, from 144·8°K upwards, a line obtained by back extrapolation of the C_p curve from 170°K upwards. The extra heat intake measured in this way was 137 ± 10 cal/mole, and in evaluating the entropy this was assumed to be absorbed at 144·8°K. The remaining, larger part of the entropy of transition was determined by graphical evaluation of $\int C_p' \, d\ln T$, where C_p' is the difference between the full and dotted curves of Fig. 5. (In estimating C_p' from 130 to 144·8°K, the full curve was assumed to be the extrapolation of that followed from 120 to 130°K.) The total entropy of transition was thus found to be 4·5 cal/mole deg., which may be in error by several tenths of a unit.

The heat capacity of rubidium stannibromide showed a slight but reproducible anomaly between 75 and 95°K (Fig. 6), which however was not large enough to complicate the heat capacity comparison with the ammonium salt.

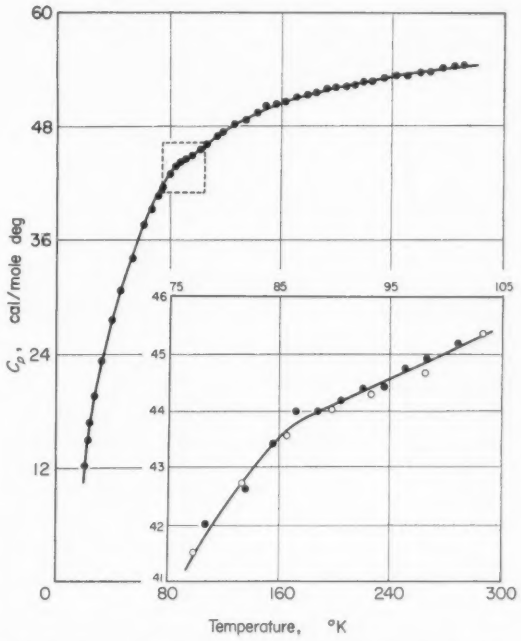


FIG. 6. C_p vs. T for rubidium stannibromide. The inset shows the region 75–105°K (two series of measurements).

DISCUSSION

We shall first discuss what conclusions can be drawn about the movement of the ammonium ion in the stannichloride and stannibromide. In Fig. 7,

$$\Delta C_p, = C_p[(NH_4)_2SnCl_6] - C_p(Rb_2SnCl_6) - -2C_{int}(NH_4^+),$$

is plotted against temperature. We must first consider how this plot may be affected by differences between the contributions from the lattice vibrations. While a reliable analysis of C_p for salts of this complexity is not possible, some guidance is provided by a comparison of the rubidium and

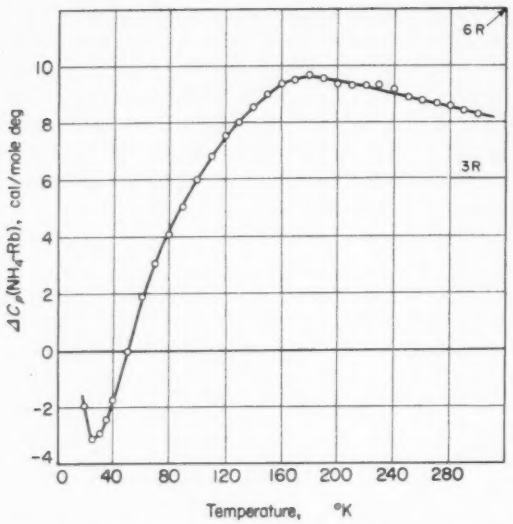


FIG. 7. $\Delta C_p(NH_4-Rb)$ for the stannichlorides vs. T. $\Delta C_p(NH_4-Rb) = C_p(NH_4$ salt$)-C_p(Rb$ salt$)-2C_{int}$ (NH_4).

potassium stannichlorides, for the difference in the unit cell dimensions of the two salts is only about 1 per cent, and so the difference in C_p at lower temperatures must be due primarily to the difference in mass of the cations. This difference is greatest at about 35°K, and by 120° (i.e. before the C_p comparison is complicated by the transition in the potassium salt, see Fig. 3) it has vanished. Presumably at about this temperature the contribution from the lattice vibrations has almost reached the limiting classical figure. It would of course be expected that for salts with such comparatively large anions the Debye temperature would be relatively low. For the ammonium and rubidium salts ΔC_p reaches its minimum at ~25°K. On any reasonable basis for the extrapolation to 0°K of the contribution from the torsional oscillations of the ammonium ions, this temperature cannot differ very much from that at which the difference between the contribution from the lattice vibrations is largest. So again, from ~120°K upwards, this last difference should be small, and in the upper region of temperature we may take the curve of Fig. 7 as giving essentially the contribution to C_p of the torsional oscillation or rotation of the two ammonium ions in the gram molecule. The failure of ΔC_p to rise asymptotically towards the limiting value of 6 R for fully excited three-dimensional simple harmonic oscillation is striking. The passage through a maximum and the subsequent decrease is in fact the behaviour expected of the heat capacity of a restricted rotator, and implies that the ammonium ions in the stannichloride are only prevented from freely rotating by a comparatively low potential barrier.

For the stannibromides with their greater unit cell sizes, the Debye temperatures should be still lower than for the corresponding stannichlorides, and the limiting contribution from the lattice vibrations to C_p should be effectively attained at still lower temperatures. The ΔC_p plot for the stannibromides (Fig. 5) is complicated by the transition in the ammonium salt, which we shall shortly discuss, but again it is significant that above the transition, i.e. from 170°K upwards, ΔC_p, instead of tending to a value of 6 R, is decreasing, and at room temperature is about midway between 3 R and 6 R. Here too the potential barrier restricting the rotation of the ammonium ion cannot be high. If the potassium salt rather than the rubidium salt is used for comparison, ΔC_p from about 240°K upwards is in fact a little less than 3 R. The rather large difference at higher temperatures between C_p for the potassium and rubidium compounds recalls the observation that potassium stannibromide, alone of these salts may have a tetragonal structure. Thus, the comparison of the ammonium with the rubidium salt would seem to be the more reliable and informative.

Entropy values for the six salts are given in Table 4. The entropy changes from 0° to 20°K were estimated by using Debye functions of the form $AD(\theta/T)$, where the coefficient A and the parameter θ were chosen to give the best fit with the experimental C_p values at 20°, 25° and 30°K. The difference in the entropies of ammonium and potassium stannichlorides at 298·16°K is 11·5 e.u., to which the internal vibrations of the ammonium ion only contribute about 0·14 e.u. Owing to the difference in Debye frequencies, the contributions from the rotational or torsional movement of the ammonium ions will, however, be larger than 11·5 e.u. by an amount which—as a comparison of the entropies of potassium and rubidium stannichlorides indicates—will probably be 2 to 3 e.u. For the stannibromides, the entropy of the ammonium salt exceeds that of the potassium and rubidium

Table 4. Entropy values, in cal/mole deg.

	20→298·16°	0→20	0→298·16°
$(NH_4)_2SnCl_6$	97·1	2·0	99·1
K_2SnCl_6	84·8	2·8	87·6
Rb_2SnCl_6	87·45	2·8	90·25
$(NH_4)_2SnBr_6$	115·7	4·5	120·2
K_2SnBr_6	101·7	4·2	105·9
Rb_2SnBr_6	102·7	3·7	106·4

VOL. 13 1960

salts by a larger amount, and here the entropy per mole of the torsional or rotational movement of the two ammonium ions might amount to about 15 e.u. The calculated rotational entropy of two gram-ions of freely rotating ammonium ions (taking 1·032 Å as the N–H bond length) is 19·70 cal/mole deg.

Since the barriers preventing rotation of the ammonium ion in both the stannichloride and stannibromide are relatively low, we have considered what light this throws on the charge distribution in the stannihalide ion by evaluating the potential energy of an ammonium ion for three different orientations with respect to its stannichloride neighbours. Only electrostatic forces have been considered, and the calculations have been confined to nearest neighbour interactions. We have supposed that each tin atom carries a charge s and each halogen atom a charge x, and that the charges are located at the centres of the atoms. Clearly $s+6x=-2$. If a is the length of the side of the unit cell, the co-ordinates of a halogen atom with respect to the tin atom in the same ion are $(\pm ua,0,0)$, $(0\pm ua,0)$ and $(0,0\pm ua)$. The value of u is $0{\cdot}240\pm0{\cdot}01$ for the stannichloride and $0{\cdot}245\pm0{\cdot}01$ for the stannibromide.[1] (If u were exactly 0·25 the centres of the halogen atoms would lie at the mid-points of the sides of the octant of the unit cell shown in Fig. 8.) The estimated potential energies are rather sensitive to the value of u, and it is unfortunate that this parameter is not known more accurately. The three orientations of the ammonium ion which we have considered are: (i) The hydrogen atoms pointing towards the tin atoms at the corners of the octant. (ii) The hydrogen atoms pointing towards the unoccupied corners of the octant. (iii) A position midway between (i) and (ii), i.e. the hydrogen

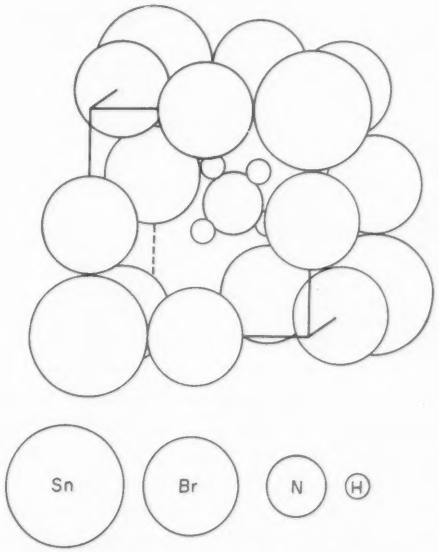


FIG. 8. An octant of the unit cell of ammonium stannibromide showing the immediate environment of the ammonium ion.

atoms pointing towards the mid-points of four of the edges of the octant. For the charge distribution in the ammonium ion we have used PLUMB and HORNIG'S estimate[10] that each hydrogen atom carries an effective charge of $+0{\cdot}43$. The results of the calculations are given in Table 5. The actual barrier heights cannot be reliably estimated from the heat capacity results, since the necessary calculations have not yet been performed for three-dimensional rotators, but from the information available on the thermodynamic properties of one-dimensional rotators it would appear unlikely that in either salt can the barrier much exceed 1000 cal. (In ammonium chloride it is 6410 cal/g

Table 5. Estimates of the relative potential energies U in cal/g ion *of an ammonium ion in ammonium stannichloride and stannibromide for the following orientations:* I, H *atoms pointing to* Sn *atoms at corners of octant;* II,H *atoms pointing to empty corners of octant;* III,H *atoms half-way between* I *and* II. *s and x are the assumed charges on the tin and halogen atoms, respectively.*

s	x	U_{III-I}	U_{II-I}
Stannichloride			
0	−0·33	1810	2910
+0·4	−0·40	390	340
+0·7	−0·45	− 670	−1580
+1	−0·5	−1740	−3510
Stannibromide			
0	−0·33	950	1520
+0·4	−0·40	− 110	− 690
+0·7	−0·45	− 910	−2340
+1	−0·5	−1710	−4000

ion, sufficiently high for the contribution from the torsional oscillations of the ammonium ion to C_p to fall well below 3 R even at room temperature.[11]) The figures of Table 5 therefore suggest that in both stannihalide ions the tin atoms carry a small positive charge, but that even in the stannichloride this does not exceed about +0·6. It also follows from Table 5 that the orientation of the ammonium ion which is preferred in the stannichloride may not necessarily be that of lowest energy in the stannibromide.

The transition in ammonium stannibromide is very extensive, covering a range of ~90°. There is a shelf in the heat capacity curve at about 120 to 135°K, such that were it not for the comparison with the rubidium or potassium salt the region of anomalous heat capacity might from the appearance of the C_p curve be supposed to be a relatively small one in the neighbourhood of 145°K. It is clear however (Fig. 5) that from about 95°K to the λ-point ΔC_p exceeds the maximum figure of 12 cal/mole deg. which can be ascribed to torsional oscillations of the ammonium ions, so that from 95°K (and probably from somewhat lower temperatures) there is an anomalous contribution to the heat capacity of the ammonium salt. The qualitative features of the transition—its extent, the shelf in the C_p curve and the hump after the maximum has been passed—suggest that it is a change of some complexity. Whereas the transition in potassium stannichloride with an entropy change of 1·13 e.u. might conceivably be due to the availability to the anions of two alternative positions, such a simple possibility cannot hold for the anomaly in ammonium stannibromide with its much larger entropy increase (~4·5 e.u.). The value of the entropy change also rules out the possibility (in any case unlikely in view of the heat capacity analysis already presented) that the anomaly involves disordering on the part of the ammonium ions. Nevertheless it is pertinent to consider what the preferred mutual orientation of a pair of adjacent ammonium ions is likely to be. As has been shown, the energy of the ammonium ion will in general be different according to whether each N–H bond is directed at a tin atom or at one of the unoccupied corners of the octant containing the ion (Fig. 8). Suppose that the former possibility is preferred. Then if any one of the six neighbouring ammonium ions adopts the same orientation with regard to the tin atoms surrounding it, it must be antiparallel to the central ammonium ion. Now two antiparallel ammonium ions, separated by a distance R, have a higher

potential energy than when in a parallel alignment at the same separation by an amount proportional to R^{-5}. Since R for ammonium stannichloride and stannibromide is 5·02 Å and 5·30 Å, respectively, this energy difference is relatively small, amounting to only about 21 per cent and 27 per cent of the corresponding quantity for ammonium chloride, for which $R = 3{\cdot}87$ Å. It is therefore likely that at lower temperatures the preferred mutual orientation of adjacent ammonium ions is an antiparallel one which is determined by the interaction of each ammonium ion with its stannihalide neighbours and with which the inter-ammonium ion interaction is not strong enough to interfere. As far as the latter interaction is concerned, any disordering of the ammonium ions with rising temperature will mean a decrease in the energy of the system.

The discovery of the λ-point in potassium stannichloride means that, at lower temperatures, the structure of this salt is less simple than that assigned to it at room temperature. It merits further investigation, but since the entropy of transition may represent $R \ln 2$ we have considered whether it is possible that there are two orientations available to each stannichloride ion which are fully utilized above the λ-point but not below. Each stannichloride ion is surrounded by eight potassium ions as shown in Fig. 9. The simple anti-fluorite structure requires that the angle $\theta = 0$. This is however not necessarily the position in which the electrostatic potential energy due to the attraction of the eight cations is a minimum. If for example the anion is rotated about the axis *ab*, this potential energy term falls to a minimum when $\theta = 45°$. On the other hand, the energy of repulsion between the metal ions and halogen atoms due to the overlap of their electronic clouds rises as θ increases from zero. A suitable combination of the energies of attraction and repulsion could therefore produce two minima in the potential energy function symmetrically disposed about $\theta = 0$. A satisfactory quantitative treatment is made difficult by uncertainty in the magnitude of the energy of replusion. Clearly the halogen atoms carrying a charge of little more than $-\frac{1}{3}$ cannot be regarded as having electron clouds as extensive as single halogen ions. We have carried out trial calculations with the repulsive potential of Huggins and Mayer[12], which necessitated choosing a value for r_-, the effective radius of the chlorine atom. We find that if this is taken to be small enough, it is in fact possible to account for a double minimum in the potential energy curve of the stannichloride ion as it is rotated about the axis *ab* in the potassium salt. Thus with $r_- = 1{\cdot}152$ Å,

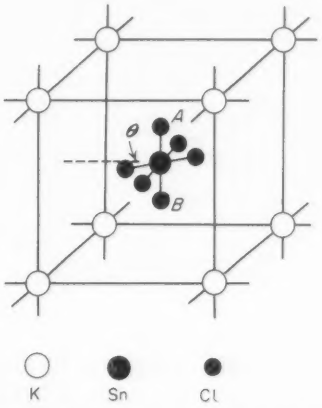


Fig. 9. An octant of the unit cell of potassium stannichloride, showing the immediate environment of a stannichloride ion.

two minima appear at $\theta = \pm 20°$ separated by about 500 cal/mole. With the same value of r_-, however, the corresponding potential energy curve for *rubidium* stannichloride (which has no λ-point) shows a minimum only at $\theta = 0^0$. If r_- is increased by the difference between the SnCl- and SnBr-bond lengths, and the calculations repeated for potassium *stannibromide* (assumed to have an anti-fluorite structure), which also has no λ-point, here too the potential energy proved to be a minimum for $\theta = 0°$. For the energy difference between two adjacent stannichloride ions at the separation prevailing in the potassium salt, according to whether (i) they are parallel with $\theta = 15°$ or (ii) one has $\theta = 15°$ and the other $\theta = -15°$,—we find 490 cal/mole, a figure of the right sort of magnitude for a λ-point at the observed temperature. On account of the sensitivity of all these estimates to the precise lattice parameters, and in view of the rather arbitrary treatment of the energy of repulsion, too much importance should not be attached to the numerical results of these calculations, for which reason they have not been treated in detail here, but the cause

VOL. 13 1960

of the λ-point in potassium stannichloride which we have discussed is at least a possibility which should be considered in the light of any further experimental information.

Acknowledgement—We thank Imperial Chemical Industries Ltd., for financial assistance.

REFERENCES

1. ENGEL G., *Z. Kristallogr.* **90**, 341 (1935).
2. KETELAAR J. A. A., RIETDIJK A. A. and VAN STAVEREN C. H., *Rec. Trav. Chim. Pays-Bas* **56**, 907 (1937).
3. MARKSTEIN G. and NOWOTNY H., *Z. Kristallogr.* **100**, 265 (1939).
4. DAVIES T. and STAVELEY L. A. K., *Trans. Faraday Soc.* **53**, 19 (1957).
5. MORFEE R. G. S. and STAVELEY L. A. K., *Nature, Lond.* **180**, 1246 (1957).
6. WAGNER E. L. and HORNIG D. F., *J. Chem. Phys.* **18**, 296 (1950).
7. SCHOCH E. P. and BROWN D. J., *J. Amer. Chem. Soc.* **38**, 1660 (1916).
8. ENGELENBURG A. J., *Z. anal. Chem.* **62**, 257 (1923).
9. FREYMANN M., *C.R. Acad. Sci., Paris* **233**, 1449 (1951).
10. PLUMB R. C. and HORNIG D. F., *J. Chem. Phys.* **23**, 947 (1955).
11. GUTOWSKY H. S., PAKE G. E. and BERSOHN R., *J. Chem. Phys.* **22**, 643 (1954).
12. HUGGINS M. L. and MAYER J. E., *J. Chem. Phys.* **1**, 643 (1933).

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 13. pp. 145–150. Printed in Great Britain.

THE ADSORPTION OF OXYGEN ON CLEAN SILICON SURFACES

MINO GREEN and K. H. MAXWELL

Zenith Radio Corporation, Chicago 39, Illinois

(*Received* 11 *June* 1959, *revised* 27 *July* 1959)

VOL. 13 1960

Abstract—The adsorption of oxygen on clean silicon surfaces produced by crushing *in vacuo* has been examined in the pressure range 30–200 μ Hg of oxygen at room temperature. The initial adsorption to the extent of about a monolayer was very rapid but kinetic data were obtained for the subsequent slow adsorption. The results were complex and probably sensitive to the crystal faces exposed during crushing. The adsorption process was found to terminate at an oxygen take-up equivalent to about 1·5 oxygen atoms per surface silicon atom. A unique configuration, with an inherent termination at a coverage of 1·5, can be constructed for oxygen adsorbed on the {111} and {110} faces of silicon. Prolonged crushing of the silicon led to an enhanced take-up of oxygen which may result from the heating of very small particles during the fast adsorption stage. The surface areas of the powders were determined by the B.E.T. method, using krypton at liquid nitrogen temperature.

INTRODUCTION AND PREVIOUS WORK

Of the two most important elemental semiconductors germanium has been studied more extensively than silicon and this situation is pronounced when the surface chemistry of the two elements is considered. The adsorption of oxygen on silicon surfaces has, however, received some attention. Law[1] has reported some results for oxygen adsorption on evaporated silicon surfaces and his main conclusions are: (i) There is an initial fast adsorption of approximately a monolayer of oxygen atoms. (ii) Beyond a monolayer there is further adsorption of oxygen which is much slower. (iii) The rate of this slow take-up can be described by the equation

$$dN/dt = a \exp(-bN) \qquad (1)$$

where N is the number of oxygen atoms adsorbed in time t, at constant temperature and pressure of oxygen, while a and b are constants. Equation (1) is often written in the equivalent form

$$N = K \log_{10} t + C \qquad (2)$$

where K and C are constants. Kafalas[2] has reported that the adsorption of oxygen on silicon follows an equation of the form of equation (2). However, the logarithmic rate law was obeyed over only a short time interval (10–80 min) and the presence of mercury cut-offs in his system means that he was observing the sum of two adsorptions which resulted in an apparent logarithmic rate law over a limited period.[3] There was an increase in $dN/d \log t$ beyond 100 min as the adsorption of oxygen on mercury became dominant. Wolsky[4], working with a single crystal of silicon, has reported the result of a single experiment in which the logarithmic rate law was apparently obeyed, but Wolsky states that his surface was not clean.

The general conclusions of Law[1] for the adsorption of oxygen on silicon have been well established for the adsorption of oxygen on germanium.[4–7] However, a closer examination of the results obtained for the germanium–oxygen system reveal some quantitative differences. In particular the values of the constants K and C (in equation 2) given by Dell[6] differ from the values reported by Green *et al.*[5] Dell[6] has demonstrated that the values of K and C are related to the mode of preparation of the germanium samples, particularly the thermal history, and concludes that the adsorption of oxygen on germanium is a "structure sensitive phenomenon". A similar

conclusion has been reached by GREEN and LIBERMAN[8].

It is to be noted that LAW[1], in his work with silicon, prepared surfaces by evaporation. Films of this kind, consisting of extremely small particles, are in a very strained state. In view of the behaviour of the germanium–oxygen system it seems likely that silicon surfaces prepared by other methods might well exhibit rather different adsorptive properties.

SCOPE OF THE PRESENT INVESTIGATION

In this work data are presented on the total take-up of oxygen and the kinetics of the slow adsorption of oxygen on powdered silicon samples produced by crushing a single crystal wafer *in vacuo*. The pressure range covered was 30–200 μ Hg of oxygen at room temperature. Surface areas were determined using the B.E.T. method.

EXPERIMENTAL TECHNIQUE

(a) *Adsorption apparatus*

The adsorption chamber was very like that described by GREEN *et al.*[5] and consisted, essentially, of two compartments separated from one another by a break-seal. In one experiment (Run No. 1) a single chamber adsorption chamber[5] was employed. This procedure involves crushing the silicon wafer in the presence of a known amount of oxygen. Typical volumes were: Silicon chamber 18 cm^3; oxygen chamber 25 cm^3; single chamber apparatus 30 cm^3.

(b) *Pressure measurement*

This was accomplished using a glass-coated thermistor bead as a non-linear Pirani gauge. The thermistor gauge circuit, while similar in principle to that described by DUSHMAN[9], has been modified in this laboratory. Calibration was against a McLeod gauge.

(c) *Surface area determination*

The B.E.T. apparatus employed was very similar to that employed by ROSENBERG[10].

(d) *Vacuum system*

The vacuum system was a conventional one and consisted of a rotary oil pump, a two-stage mercury diffusion pump and a trap kept at liquid nitrogen temperature. After baking out the adsorption apparatus at 300°C for 4 hr, a vacuum of better than 10^{-6} mm Hg could be readily attained. On standing for 60 hr the pressure rose to only 3×10^{-6} mm Hg.

(e) *Materials*

The silicon discs employed in this work were cut from a single crystal of *n*-type silicon ($\rho = 12$–$25\ \Omega$-cm), were approximately 1 cm in diameter and 0·04 cm thick. Both the oxygen and krypton were reagent grade gases.

(f) *Experimental procedure*

The apparatus was evacuated and baked at 300°C for several hours. The lower chamber was then sealed off. A known quantity of oxygen was admitted to the upper chamber which was then sealed off. The silicon disc in the lower chamber was crushed using a glass-encased drop-hammer. To start an experimental run the break-seal, connecting the two chambers, was broken with another drop-hammer and the pressure recorded as a function of time. The volumes of both the lower and upper compartments were measured before each run. At the conclusion of the experiment the surface area of the powder was determined.

(g) *Accuracy*

The total take-up of oxygen could be determined with an accuracy of ±2 per cent. The surface area determinations were reproducible to about ±3 per cent. The absolute accuracy of the B.E.T. method is a controversial point,[10–11] but a reasonable estimate would seem to be ±10 per cent.

(h) *Cleanness of the silicon surfaces*

Cleanness of the adsorbent surface is vital in all adsorption studies and a discussion of this matter follows. First, it must be emphasized that we are dealing with a large adsorbent surface area ($\sim$300 cm^2) in a sealed, all-glass system of comparatively small volume ($\sim$40 cm^3). This is to be contrasted with many adsorption systems used in semiconductor surface work where the surface area may be small (1–10 cm^2) and the volume of the system large ($\sim$1000 cm^3).

The glass of the adsorption chamber is neither a source nor a sink of gas to any significant extent. When the apparatus stood, sealed in vacuum (with

VOL. 13 1960

the silicon uncrushed), the pressure rose to only 3×10^{-6} mm Hg after 60 hr. The amount of gas desorbed from the walls of the apparatus in this time was, therefore, approximately 0·15 μ cm^3. To cover 300 cm^2 of silicon surface with a monolayer of oxygen atoms requires approximately 5000 μ cm^3. Obviously desorption of chemisorbable material from the walls can contaminate only a minute fraction of the total surface area during the course of an experimental run. A blank double chamber run was carried out with no silicon present and it was found that the pressure of oxygen in the system after cracking the break-seal was that expected for zero adsorption. Therefore the glass walls do not adsorb oxygen to within the limits of error of our experiments.

Another possible source of contamination is impurities contained in the bulk of the silicon itself. It is known that many gases dissolve in silicon (and germanium) and, since crushing in part liberates them, they may contaminate the adsorbent surface. The following experiment was carried out. Silicon was crushed in vacuum and the pressure followed during the crushing operation. No pressure rise was noted and hence no measurable amount of non-adsorbable gas was given off from the silicon. This is in contrast to the results obtained with germanium grown in argon.[5] The only other source of surface contaminant is by diffusion from the bulk of the powder particles. That both the single and double chamber experiments gave essentially the same results is good evidence that the diffusion process did not produce significant contamination. In the single chamber runs the crushing was carried out in the presence of the oxygen and the rate of adsorption of oxygen on silicon must be much greater than the rate at which any contaminant could diffuse from the interior of the powder particles. In the double chamber technique the silicon was crushed in vacuum and the oxygen admitted some time later.

RESULTS

The results of several runs are given in Fig. 1 in which the coverage (i.e. the average number of oxygen atoms per surface silicon atom) is plotted against the logarithm of the time. Fig. 1 also lists the oxygen pressure in the adsorption chamber at the end of the reaction. The coverage at a particular time was calculated from the oxygen take-up and the surface area, assuming a figure for the number of adsorption sites per cm^2. The number of broken valence bonds and the number of surface atoms per cm^2 for the three principal crystal faces of silicon are:

{111} planes $7{\cdot}84\times10^{14}$ sites and bonds
{110} planes $9{\cdot}58\times10^{14}$ sites and bonds
{100} planes $6{\cdot}78\times10^{14}$ sites and $13{\cdot}6\times10^{14}$ bonds.

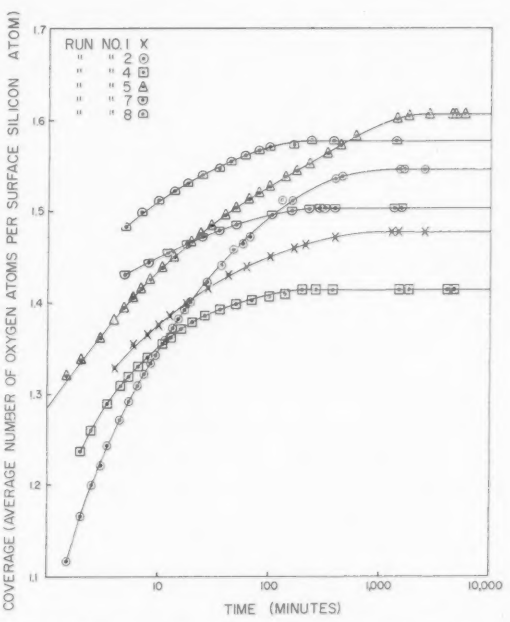


FIG. 1. The adsorption rate of oxygen on clean silicon surfaces. Run No. 1, single chamber apparatus (final pressure 70μ Hg, surface area 210 cm^2); Run No. 2 (200μ Hg, 319 cm^2); Run No. 4 (26μ Hg, 331 cm^2); Run No. 5 (99μ Hg, 233 cm^2); Run No. 7 (178μ Hg, 299 cm^2); Run No. 8 (160μ Hg, 217 cm^2).

It seems unlikely that an appreciable amount of the {100} faces could be formed in competition with the other two, since twice the number of Si–Si bonds per surface atom must be broken to expose this face. Consequently we have selected $8{\cdot}7\times10^{14}$ sites/cm^2, the mean of the values for the {111} and {110} planes, as a reasonable figure. In so doing, we have assumed that these two faces, in approximately equal amounts, formed the majority of the surface produced by crushing silicon wafers. The true distribution would have to be drastically

different to alter this figure by more than a few per cent.

DISCUSSION

Several features of the results in Fig. 1 merit attention. All the experimental runs showed a very definite termination in the take-up of oxygen by the silicon surface. Beyond this point no further adsorption of oxygen could be detected. It can be seen that the values of the final coverage reached were in the range 1·4–1·6. This spread of $\pm 0{\cdot}1$ coverage units is within the error imposed by the surface area and other determinations. In view of the additional uncertainty resulting from the choice of a value for the number of adsorption sites per cm^2, a reasonable conclusion seems to be that the final coverage is 1·5 oxygen atoms per surface silicon atom, with an error of $\pm 0{\cdot}2$ coverage units.

It should be noted that the logarithmic rate law (equation 2) is not obeyed in these cases; none of the experimental runs show a portion where the uptake of oxygen increases linearly with the logarithm of the time. These results are then to be contrasted with those obtained for the adsorption of oxygen on germanium[5] where the logarithmic rate law was obeyed over the whole time of measurement (6 sec–22 days), during which period the coverage increased from 0·9 to 1·9.

It is of great interest that a unique configuration, giving an inherent termination at a coverage of 1·5, can be constructed for oxygen adsorbed on the {111} and {110} faces of silicon. The relative dispositions of oxygen and silicon are somewhat different on the two planes but in both cases an extended array can be built up. The configuration is obtained assuming that the normal covalencies of silicon and oxygen must be satisfied. While models of these structures have been built, the principle is best illustrated diagrammatically and the steps are imagined to be as shown in Fig. 2. The inversion of alternate O–Si groups, shown in step 1 of the diagram, occurs under the influence of the repulsion between neighbors. The problem then is to satisfy the valencies of the three underlying silicon atoms previously attached to the silicon atom now on the surface. This can be done if it is assumed that one of these silicon atoms bonds to the oxygen atom now in the second layer, while the other two silicon atoms bond to one another.* The silicon atom, now on top, bonds to the oxygen atom attached to the neighboring surface silicon atom forming a cyclic structure. Ring structures containing alternate oxygen and silicon atoms are well known as the basic unit in many silicate minerals.[15, 16] We visualize the last stage

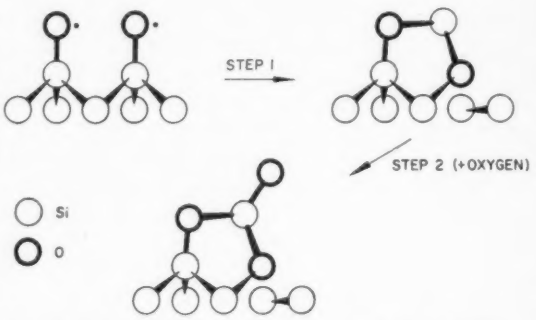


FIG. 2. Suggested steps in the formation of a configuration corresponding to a coverage of 1·5. The starting point is the silicon surface covered with a monolayer of oxygen.

to be the reaction of oxygen with the still unsaturated silicon in the top layer, giving a fully saturated structure corresponding to a coverage of 1·5. This structure is probably metastable, due to the comparatively weak underlying Si–Si bond, and can be broken down (by raising the temperature for example), thereby allowing the process of oxide formation to occur. According to this picture of the process the adsorbed phase corresponding to a coverage of 1·5 is essentially covalent. It is well known that SiO_2 is a largely ionic solid and so, for take-up of oxygen beyond 1·5 coverage to occur, there must be a change in mechanism. Further uptake would probably occur via the diffusion of silicon ions through to the surface or by the diffusion of oxygen ions into the solid.

Another consequence of our mechanism is that we would expect the adsorption of oxygen on the {100} faces to be very different. The {100} faces of silicon (and germanium) have two "free valencies"

* It is to be noted that FARNSWORTH and his associates,[12–14] in their slow speed electron diffraction studies of cleaned germanium and silicon surfaces, have obtained results consistent with the existence of a surface bond. While this is not the same as the bond between underlying silicon atoms, such as we are postulating, there are obvious similarities.

VOL. 13 1960

per surface atom and a valence saturated structure can be constructed only for a coverage of 1·0 oxygen atoms per surface silicon atom. Beyond that point further uptake can occur only by what we can call oxide formation, as distinct from chemisorption, with consequently different kinetics.

The markedly different shapes of the curves shown in Fig. 1 pose a problem. It can be seen that the pressure of oxygen has no obvious effect on either the final coverage or the shape of the curves. While the errors in the B.E.T. and other determinations can move these curves relative to one another, this by itself is not a sufficient explanation. There remained the possibility that the oxidation of some fraction of the surface may have been diffusion controlled with the rate determining step being the diffusion of oxygen along fine capillaries in the powder. This was ruled out because agitating the powder, using an externally applied vibrator, caused no sudden break in the plot of oxygen take-up versus time. It was therefore concluded that a diffusion process was not rate controlling at any time during the measured period of oxygen take-up.

We consider the most likely cause is that the crushing of a silicon wafer is not reproducible in the sense that somewhat different proportions of the two crystal faces were exposed in each experimental run. There is reason to think that the rates of adsorption will be different on different crystal faces. In the same way it can be argued that small and varying amounts of other crystal faces were produced in each crushing. With the data we have, it is not possible to distinguish between the two possibilities; perhaps both operate. We believe that the explanation for the variety of shapes exhibited by the curves in Fig. 1 is a structural one such as we have mentioned and that we were observing the sum of at least two adsorptions proceeding at rather different rates.

Some further data were obtained in the following manner. After the surface area measurement, the tube and powder were resealed on the adsorption apparatus and pumped long enough to ensure a good vacuum. The powder was crushed further, a dose of oxygen admitted and another take-up of oxygen measured. The total area of this powder was determined at the end of the experiment and the fresh surface area produced in the second crushing was calculated by difference. This procedure was carried out for the powders labeled (7) and (8) in Fig. 1 and the results of the second crushings are given as curves (7b) and (8b) in Fig. 3, in which again the coverage is plotted against the logarithm of the time. It is important to note that the assumption is made that no desorption of oxygen occurs from that part of the total surface already covered by oxygen; this was found to be so by LAW[1].

The most important feature of the two runs shown in Fig. 3 is the greatly enhanced take-up of

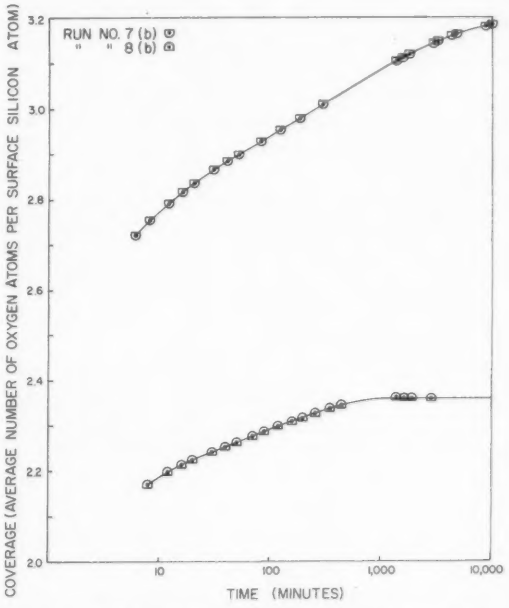


FIG. 3. The adsorption rate of oxygen on clean silicon surfaces, second crushing. Run No. 7(b) (final pressure 51 μ Hg, surface area 231 cm^2); Run No. 8(b) (111 μ Hg, 234 cm^2).

oxygen, the final coverage having the values of 3·2 and 2·4 respectively, for Runs (7b) and (8b). Also, it should be noted that the take-up of oxygen had already exceeded a coverage of 2 before the first reading was taken.

It seems to us that there are two possible explanations for the results with the finer powders.

(i) The extensive crushing may have exposed crystal faces on which the configuration corresponding to a coverage of 1·5 either cannot form—the (100) faces for example—or is no longer stable.

(ii) The heat of adsorption on very small particles, which have a large surface/volume ratio, can raise their temperature substantially leading to an enhanced take-up of oxygen and possibly a different mechanism. Such an effect has been observed by STONE and his collaborators[17,18] for the adsorption of oxygen on both nickel powders and, to a lesser extent, on evaporated nickel films. It is obvious that the two effects (i) and (ii) are not mutually exclusive and that both may operate simultaneously.

Finally, it should be said that we regard these results for the adsorption of oxygen on the finely divided powders as being, in a sense, anomalous. The results for coarse powders (Fig. 1) represent the normal mode of adsorption at room temperature, such as would be expected for single crystals. In view of the complexity of the results further work on this problem would best be carried out using well defined single crystal surfaces.

Acknowledgement—The authors wish to thank Dr. J. T. LAW of Bell Telephone Laboratories Inc. for providing the silicon used in this work.

REFERENCES

1. LAW J. T., *J. Phys. Chem. Solids* **4**, 91 (1958).
2. KAFALAS J. A., *Quart. Progr. Rep. Lincoln Lab.* p. 19, Nov. 1955.
3. KAFALAS, J. A., private communication.
4. WOLSKY S. P., *J. Phys. Chem. Solids* **8**, 114 (1959).
5. GREEN M., KAFALAS J. A. and ROBINSON P. H., *Semiconductor Surface Physics* (Edited by R. H. KINGSTON) p. 349. University of Pennsylvania Press, Philadelphia (1957).
6. DELL R. M., *J. Phys. Chem.* **61**, 1584 (1957).
7. SPARNAAY M. J. in *Proc. Internat. Conf. Solid State Physics and its Applic. to Electronics and Telecomm., Brussels, June* 2–7, 1958. To be published by Academic Press, New York.
8. GREEN M. and LIBERMAN A., to be published.
9. DUSHMAN S., *Scientific Foundations of Vacuum Technique* p. 330. J. Wiley, New York (1949).
10. ROSENBERG A. J., *J. Amer. Chem. Soc.* **78**, 2929 (1956).
11. JACOBS P. W. M. and TOMPKINS F. C., *Chemistry of the Solid State* (Edited by W. E. GARNER) p. 91. Academic Press, New York (1955).
12. SCHLIER R. E. and FARNSWORTH H. E., *Semiconductor Surface Physics* (edited by R. H. KINGSTON) p. 3. University of Pennsylvania Press, Philadelphia (1957).
13. FARNSWORTH H. E., SCHLIER R. E., GEORGE T. H. and BURGER R. M., *J. Appl. Phys.* **29**, 1150 (1958).
14. FARNSWORTH H. E., SCHLIER R. E. and DILLON J. A., *J. Phys. Chem. Solids* **8**, 116 (1959).
15. WELLS A. F., *Structural Inorganic Chemistry* (2nd Edition. p. 569) Oxford University Press (1950).
16. EVANS R. C., *An Introduction to Crystal Chemistry* p. 223. Cambridge University Press (1939).
17. DELL R. M., KLEMPERER D. F. and STONE F. S., *J. Phys. Chem.* **60**, 1586 (1956).
18. KLEMPERER D. F. and STONE F. S., *Proc. Roy. Soc.* A**243**, 375 (1958).

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 13. pp. 151–153. Printed in Great Britain.

SOLID SOLUTION IN $A^{II}B^{VI}$ TELLURIDES

J. C. WOOLLEY and B. RAY

Department of Physics, University of Nottingham, England

(*Received* 9 *September* 1959; *revised* 28 *October* 1959)

Abstract—Alloys have been produced for the three systems CdTe–HgTe, CdTe–ZnTe and HgTe–ZnTe and annealed to obtain equilibrium conditions. It has been confirmed that solid solution occurs at all compositions in each system and the variation of lattice parameter with composition has been determined in each case. The form of the solidus curve has been obtained by X-ray methods in the HgTe–ZnTe system.

VOL.
13
1960

1. INTRODUCTION

INTEREST has increased recently in the alloy semiconductors obtained when solid solution occurs between two intermetallic compounds. Early work on $A^{II}B^{VI}$ tellurides showed some ranges of solid solution and more recently various workers have obtained quantitative data on the effects of solid solution with these compounds. GORYUNOVA and FEDOROVA[1] showed that considerable solid solution occurred in the CdTe–ZnTe system. Recently KOLOMIETS and MALKOVA[2] made approximate measurements of the variation of lattice parameter with composition for this system and investigated the electrical and optical properties of the alloys. NIKOLSKAYA and REGEL[3,4] made similar investigations on the HgSe–HgTe system. Very recently, LAWSON *et al.*[5] have made electrical and optical measurements on the CdTe–HgTe system and determined the variation of certain semiconductor parameters as a function of composition.

In work on alloy semiconductors, it is of considerable interest to know the variation of lattice parameters with composition, for then the determination of these parameters for an alloy gives, within certain limits, a relatively accurate and rapid determination of its composition. This is of particular importance when alloys are subjected to such techniques as directional freezing and zone melting, which give a variation of composition along the length of the ingot.[6,7] Of the available $A^{II}B^{VI}$ telluride systems having zinc blende structure, viz. CdTe–ZnTe, CdTe–HgTe and HgTe–ZnTe, the only available lattice parameter values are the approximate values of KOLOMIETS and MALKOVA. As an initial part of a programme of work involving $A^{II}B^{VI}$ alloys, it was decided to determine the lattice parameter variations in the three systems mentioned above, and these results are presented here. A further useful quantity in alloy semiconductor work is the solidus curve for the system concerned. Here the possibility of determining the solidus by X-ray techniques[6] has been considered.

2. PREPARATION AND METHOD

The compounds concerned were each prepared by synthesis from the elements in vacuum in sealed quartz tubes, and then each of the alloys was prepared from the appropriate compounds. The elements used were initially purified by distillation, starting from commercially pure materials.

The alloys in the systems CdTe–ZnTe and CdTe–HgTe were prepared by melting the appropriate compounds together under vacuum and then annealing the alloys until X-ray photographs indicated that equilibrium had been attained; the annealing temperature always being below the melting point of the lower melting point component. Satisfactory results were obtained with the CdTe–HgTe system after 2 days annealing at 600°C and with the CdTe–ZnTe system after 7 days at 970°C. In the case of the HgTe–ZnTe alloys, it was found difficult to produce alloys by melting because the molten ZnTe tended to attack the quartz which could not then withstand the pressures produced by the HgTe. These alloys

were therefore produced by thoroughly mixing the compounds in powder form, compressing and then annealing the powder under vacuum. By this method, good equilibrium conditions could be obtained for these alloys in a relatively short time. Thus alloys of low Hg content required to be annealed for 14 days at 600°C while the alloys with high Hg content required 28 days at the same temperature.

To minimize loss of any of the components, which could be a particular problem of this type of investigation, the alloys were always sealed under vacuum in small quartz tube before any heating, and proper annealing periods were determined by means of several trial runs. Chemical analyses of specimen alloys showed no deviations from weighed out compositions of greater than 2 mol. per cent. Such analyses do not, of course, give any indication of whether dissociated material is present.

In work on solid solution in mixed $A^{III}B^{V}$ systems,[6] it has been shown that the solidus curve of the system can be obtained by X-ray methods in the case of alloys for which low diffusion rates cause a slow approach to equilibrium. Each point in the solidus curve is obtained by annealing a suitable alloy at a temperature corresponding to a point in the liquid–solid two phase field, and then rapidly quenching to retain the equilibrium solid phase. The composition of this solid phase, determined from its lattice parameter, gives the point on the solidus curve at the annealing temperature. This technique was attempted for the three systems discussed here, but was successful in only one of these, viz. the HgTe–ZnTe system. In the case of the CdTe–ZnTe system, the diffusion rates in the solid state were too rapid for the method to be of any value, while in the case of the CdTe–HgTe system, the variation of lattice parameter from one compound to the other is so small that no solidus points determinations would be of any significance even if the diffusion rates were suitable.

The possible existence of closed miscibility gaps was also investigated. This was done by bringing the alloys to a single phase state and then annealing for a suitable time at an appropriate lower temperature. Any break up of the single phase condition would indicate the presence of a closed miscibility gap.

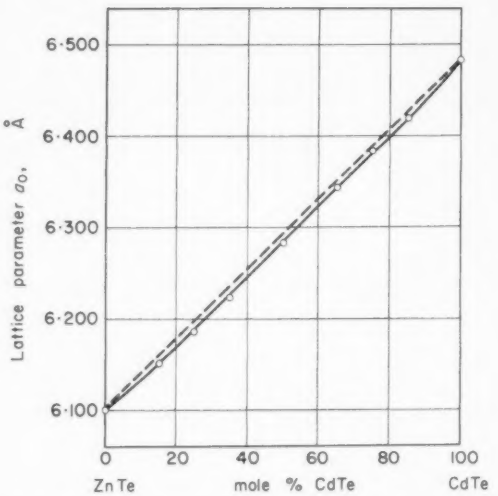


FIG. 1. Variation of lattice parameter a_0 as a function of composition in the system CdTe–ZnTe.

3. RESULTS

In all three systems, solid solution was obtained throughout the whole range of composition, the alloys having the usual zinc blende structure. The variation of lattice parameter with composition for the three systems is shown in Figs. 1, 2 and 3.

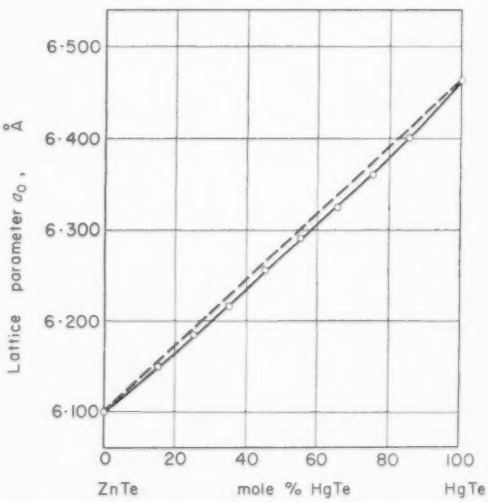


FIG. 2. Variation of lattice parameter a_0 as a function of composition in the system HgTe–ZnTe.

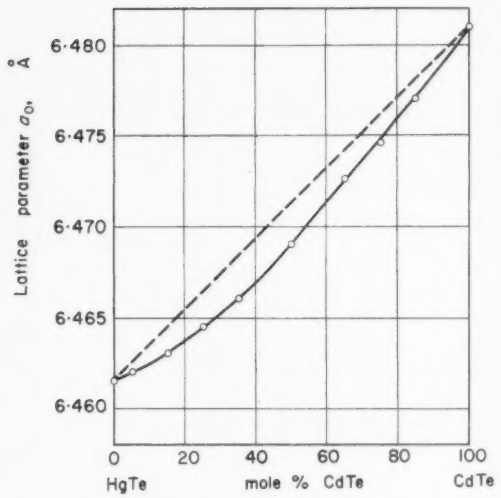


FIG. 3. Variation of lattice parameter a_0 as a function of composition in the system CdTe–HgTe.

The results for the systems CdTe–ZnTe and HgTe–ZnTe, where a large variation in lattice parameter occurs between the compounds, were obtained by using a 9-cm Unicam powder camera and CuKα radiation. In these cases (Figs. 1 and 2) the accuracy of lattice parameter is estimated as ± 0·002 Å. In the CdTe–HgTe system the variation of lattice parameter with composition is small, and so a precalibrated 19-cm Unicam powder camera was used. In this case the accuracy of lattice parameter is estimated as 0·0005 Å. It is seen that in all cases the deviation from the Vegard line is small, having a maximum value of approximately 0·002 Å for the CdTe–HgTe system, 0·004 Å for the CdTe–ZnTe system, and 0·006 Å for the HgTe–ZnTe system.

In all three systems, no sign of a miscibility gap was observed at temperatures down to 500°C, nor was any sign of ordering observed in the systems.

The solidus curve determined for the HgTe–ZnTe alloys is shown in Fig. 4, where the accuracy of temperature is estimated as ±5°C and the accuracy of composition as ±3 mol. per cent. As a check on the accuracy of the method in this case, alloys of different composition were held at the same temperature and a solidus point determined from each. The agreement between results obtained in this way was within the limits quoted above.

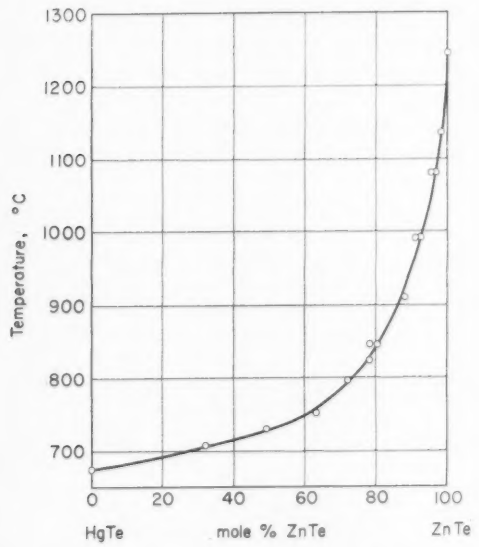


FIG. 4. Solidus curve for system HgTe–ZnTe.

Acknowledgements—We are indebted to Dr. M. E. HAINE of Messrs. Siemens Edison Swan Ltd. for a maintenance grant to one of us (B.R.). We also wish to thank Mr. F. G. STICKLAND of Messrs. Siemens Edison Swan Ltd. for the chemical analysis of specimens.

REFERENCES

1. GORYUNOVA N. A. and FEDOROVA N. N., D.A.N. **40**, (6), 1039 (1953).
2. KOLOMIETS B. T. and MALKOVA A. A., *J. Tech. Phys., Moscow* **28**, 1662 (1958).
3. NIKOLSKAYA E. J. and REGEL A. R., *J. Tech. Phys. Moscow* **25**, 1347 (1955).
4. NIKOLSKAYA E. J. and REGEL A. R., *J. Tech. Phys. Moscow* **25**, 1352 (1955).
5. LAWSON W. D., NIELSON S., PUTLEY E. H. and YOUNG A. S., *J. Phys. Chem. Solids* **9**, 325 (1959).
6. WOOLLEY J. C. and SMITH B. A., *Proc. Phys. Soc. Lond.* **72**, 214 (1958).
7. WOOLLEY J. C., EVANS J. A. and GILLETT C. M., *Proc. Phys. Soc. Lond.* **74**, 244 (1959).

VOL. 13 1960

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 13. pp. 154–155. Printed in Great Britain.

PHOTOCONDUCTION IN TERNARY V–VI–VII COMPOUNDS

R. NITSCHE and W. J. MERZ

Laboratories RCA Ltd., Zurich, Switzerland

(*Received* 31 *August* 1959)

Abstract—Some photoelectric properties of single crystals of ternary compounds of the type V–VI–VII (V = Sb, Bi; VI = S, Se, Te; VII = Cl, Br, I) were studied. The wavelengths of maximum photocurrent, λ_{max} shift in a regular way towards longer wavelengths with increasing atomic weight of the components. For the compounds SbSBr and SbSI the dependence of photocurrent, dark current and λ_{max} on temperature was measured between −140 and +120°C.

THE formation of V–VI–VII compounds (where V = Sb, Bi; VI = S, Se, Te; VII = Cl, Br, I) has been observed already in the past century.[1, 2]

In 1950, E. DÖNGES[3] studied the structure of these compounds by X-ray analysis. He found that most of the members of the V–VI–VII family belong to the rhombic-bipyramidal system, space group D_{2h}^{16}. From his data on colour and variation of colour with composition it could be assumed that these compounds have band gaps in the visible and the near infrared of the spectrum. It therefore seemed interesting to prepare single crystals of these materials and study their photoelectric properties as a function of composition. The compounds form readily by reacting equivalent amounts of the elements between 500 and 600°C in sealed, evacuated glass ampoules. To avoid the use of gaseous halogens it is also possible to start with a mixture of a group-V-trihalide and the corresponding group-V-trichalcogenide, e.g. $SbBr_3 + Sb_2S_3 \rightleftarrows 3SbSBr$.

As indicated by the equation, the V–VI–VII compounds do not melt congruently but undergo decomposition into the trichalcogenide and the trihalide. To suppress this decomposition—which is more pronounced the higher the vapour pressure of the trihalide—a surplus of the trihalide was added to the reaction mixture.

Single crystals in the form of characteristic needles, up to 2 cm in length, were obtained by slowly cooling the molten compound at a rate of 5°C/hr. The crystals were washed with diluted hydrochloric acid to remove any surplus trihalide, rinsed with alcohol and dried. They are stable in air but are attacked by strong acids. The needles show a remarkable degree of plasticity. They can be bent easily and cleave along the needle axis.

Contact was made to the crystals by gallium or silver paste. The latter, however, gave less reproducible results because at high fields after a few minutes a drop in photocurrent was noted, probably due to an electrolytic action between contacts and the crystal (formation of silver halide).

The following table shows the compounds prepared and the wavelength, of maximum photocurrent, λ_{max}

Material	Colour	λ_{max}:
SbSBr	orange needles	5450–5500 Å;
SbSI	red needles	6300–6400 Å;
SbSeBr	dark red needles	6500 Å;
SbSeI	black needles	6500 Å;
SbTeI	black needles	7200 Å;
BiSCl	red needles	5900 Å;
BiSBr	dark red needles	6350 Å;
BiSI	black needles	7850 Å
BiSeCl	dark red needles	high dark current
BiSeBr	black needles	high dark current
BiSeI	black needles	high dark current

λ_{max} seems to depend also on the composition of the crystal. Crystals of SbSI obtained from different batches of varying stoichiometric ratios between Sb_2S_3 and SbI_3 showed differences in λ_{max} ranging from 5800 to 6400 Å.

The most sensitive crystals obtained so far are SbSI and SbSBr. Both, the photocurrent I_P and the dark current I_D show an ohmic behaviour for Ga and silver paste electrodes. No saturation voltage has been observed yet. At a light flux of approximately 10^{15} photons/sec cm^2 ratios of a few thousands were obtained for I_P/I_D at λ_{max}.

The sensitivities of SbSBr and SbSI at room temperature are of the order of 10^{-4} mho cm^2/W, which is comparable to the sensitivity of HgI_2 crystals as reported by BUBE[4]. It is expected that these values can be increased by further purification or proper doping of the crystals. Assuming a mobility of $\mu = 1$ cm^2/V sec we obtain a lifetime of the order of $1\text{–}5 \cdot 10^{-5}$ sec at room temperature.

Most samples of SbSI show on a log–log scale a linear dependence of the photocurrent I_P on light intensity L. The value of the exponent n in the relation $I_P \sim L^n$ is between 0·4 and 0·6 over a range of four orders of magnitude in L.

The photocurrent in SbSBr and SbSI is quite temperature dependent. The dark current stays fairly constant between -140 and $+100$°C. The photocurrent I_P increases roughly exponentially with temperature. At -140° I_P is barely measurable. At 100° C the sensitivity has increased by at least an order of magnitude reaching values around 10^{-3} mho cm^2/W. Above 100°C I_D starts to increase and at $+150$°C the crystals begin to decompose.

The temperature variation of the bandgap E as derived from the temperature dependence of λ_{max} obeys the usual law: $E = E_0 - \beta T$. The values of the coefficients β are remarkably high. We found for SbSBr: $\beta \cong 9 \times 10^{-4}$ eV/degree

SbSI : $\beta \cong 15\text{–}16 \times 10^{-4}$ eV/degree.

We wish to thank Mr. H. U. BOELSTERLI for his help in growing the crystals.

REFERENCES

1. SCHNEIDER R., *Ann. Phys. Lpz.* **93**, 466 (1854). *Ann. Phys. Lpz.* **94**, 630 (1855); *Ann. Phys. Lpz* **110**, 150 (1860).
2. OUVRARD L., *C.R. Acad. Sci., Paris* 1516 (1893). *Ann. Chim. Phys.* **7**, II, 234 (1894).
3. DÖNGES E., *Z. Anorg. Chem.* **263**, 112 (1950); *Z. Anorg. Chem.* **263**, 280 (1950); *Z. Anorg. Chem.* **265**, 50 (1951).
4. BUBE R. H., *Phys. Rev.* **106**, 703 (1957).

VOL. 13 1960

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 13. pp. 156–159. Printed in Great Britain.

FURTHER EXPERIMENTS CONCERNING THE SPIN-ELECTRON INTERACTIONS IN SUPERCONDUCTORS

B. T. MATTHIAS, H. SUHL and E. CORENZWIT

Bell Telephone Laboratories, Inc., Murray Hill, New Jersey

(*Received* 14 *September* 1959)

Abstract—The lowering of the superconducting transition temperature of lanthanum by dissolving in it the rare earths Ce, Nd, Sm, Eu, Tb, Ho and Er is reported. The discrepancies regarding previous theoretical considerations are pointed out.

IN an earlier note[1] we have shown that the superconducting transition temperature, T_c, of lanthanum is lowered rather sharply if small amounts of rare-earth elements are dissolved in it. The depression was essentially a function of the projection of the spin on the lowest J state only, although europium is an exception to this rule.[2] The solid solutions used to illustrate this rule were those of a 1 atomic per cent rare-earth solution in lanthanum, and deviations from the theory were attributed to scatter. The variation of T_c with the concentration of the rare-earth impurity was demonstrated by a continuous plot showing T_c vs. the concentration of gadolinium dissolved in lanthanum.

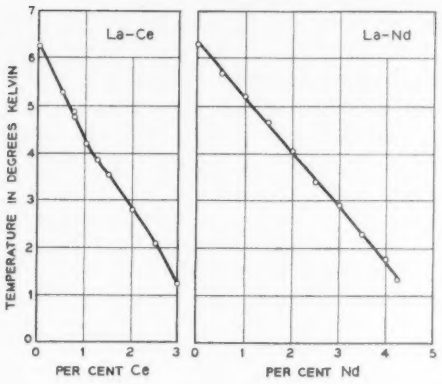
The purpose of the present paper is to give these continuous plots for several rare-earth impurities dissolved in lanthanum (Figs. 1 (a)–(g)).

These data have two surprising features. One is that cerium does not act as the above ideas would suggest. It lowers T_c by amounts comparable to elements with larger electronic spin, which may be due to the well-known unstable situation of its $4f$ electron. The other is that neodymium lowers T_c by about twice the amount produced by a corresponding concentration of erbium. While we formerly concluded that the depression was essentially symmetrical with respect to gadolinium, the systematic deviation from this behavior becomes quite apparent. Both neodymium and erbium are in symmetrically equivalent positions. They have the same absolute spin and orbital moment and much the same projection of S on the lowest J-state. Due to spin-orbit coupling, the effective moment of erbium is, however, much larger than that of neodymium. It is therefore rather surprising that the latter depresses the superconducting transition in lanthanum much faster than does erbium. If one considers secondary factors like volume and mass, this effect is even more pronounced, as according to them erbium should lower the critical temperature even more.

The question arises as to why a larger effective moment should favor superconductivity, a phenomenon that has been observed before repeatedly. By dissolving the magnetic transition elements from chromium to cobalt in titanium, we have shown that the superconducting transition temperature of titanium is raised considerably more than might be expected from a variation of valence-electron concentration.[3] Apart from stressing this analogy and implying that a weak magnetic interaction raises the superconducting transition, we have at present no other explanation for the difference between neodymium and erbium.

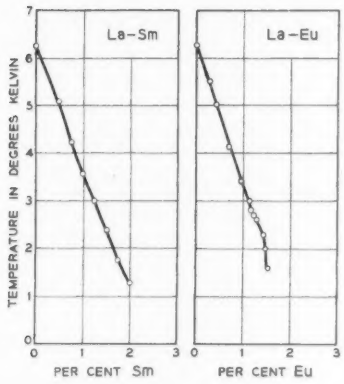
The metal samples were prepared by fusion of the elements and subsequent quenching in an argon arc furnace. The weight loss was negligible ($<0{\cdot}03$ per cent) and no attempt was made to analyze the samples. The weight loss in europium was somewhat larger ($<0{\cdot}12$ per cent) and was attributed entirely to evaporation. The melts were rapidly quenched, as is characteristic of the arc-furnace method. They were then cut with some care to minimize mechanical strains and formed

VOL. 13 1960



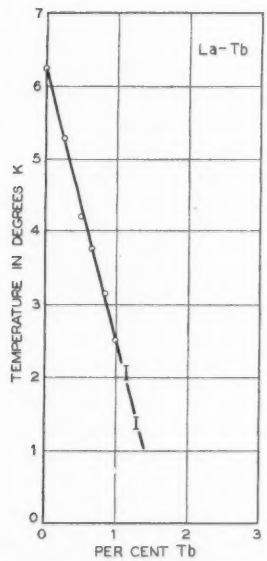
(a)

(b)

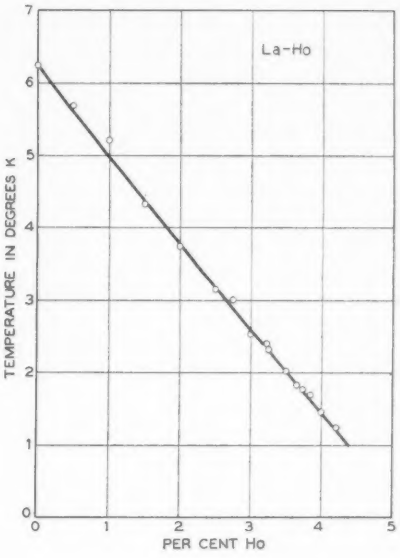


(c)

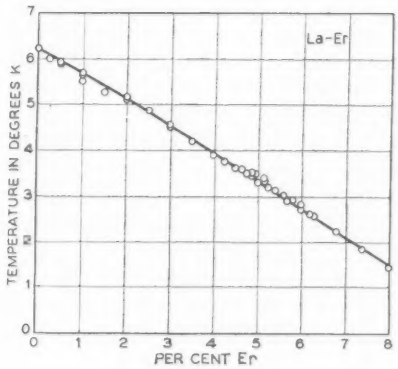
(d)



(e)



(f)



(g)

FIGS. 1. (a)–(g). Superconducting transition temperatures of lanthanum–rare-earth solid solutions.

the core of a mutual inductance for the measurement of the superconducting transition. Superconductivity was observed by determining the volume susceptibility as previously described.*[4] While this method encountered no difficulties with the rare earths from cerium to gadolinium, the elements following gadolinium did not seem to give very reproducible data at first. This was due to the higher melting points of terbium, holmium and erbium, a difficulty which was overcome by first preparing master alloys that consisted of 75 atomic per cent rare earth and 25 atomic per cent lanthanum. With these the final compositions in lanthanum were then prepared.

The superconducting transition temperatures were fairly sharp in most cases, with the exception of terbium solid solutions.

In the course of trying to improve the sharpness of the transitions in these terbium solid solutions, we found that terbium precipitated out almost entirely in the course of a 5-day anneal at 650°C, and the transition temperature reverted to that of pure lanthanum. This situation was reversible, i.e. after remelting and quenching, the original behavior was restored. This was rather unexpected. Annealing now the solid solutions of neodymium, samarium and gadolinium in the same way, we observed this precipitation in varying degrees in all cases. A lanthanum melt containing 3·5 atomic per cent neodymium thus lost 1·5 atomic per cent neodymium. In view of the strict isomorphism of lanthanum and neodymium and the small difference (3 per cent) between their lattice constants, complete miscibility at almost any temperature would have been expected. Whether this precipitation is due to magnetic effects can only be conjectured at this time.

It remains to discuss how these data fit the theory[2, 5] based on exchange scattering. We recall that the theoretical evidence strongly favored the view that exchange scattering depressed the transition temperature, not so much by reducing the energy gap of the superconducting state, but rather by reducing the free-energy difference of the normal and superconducting states. As a result the transition takes place at a finite energy gap, i.e. it should be of first order. This theory also predicted that the curve of transition temperature T_c vs. impurity concentrations ξ should be convex, beginning with $dT_c/d\xi = 0$ at $\xi = 0$ and ending with $dT_c/d\xi = \infty$ at $T_c = 0$. This theoretical result is in some doubt only near $\xi = 0$, where the small change that does occur in the energy gap dominates the picture and gives some small initial slope to the T_c vs. ξ curve. Previous experimental evidence seemed in good accord with this prediction, in particular there were strong indications of a vertical tangent near $T_c = 0$.[6]

The present data do not (except perhaps in the case of lanthanum–europium) reach down far enough in temperature to confirm the vertical tangent. But quite aside from this, the remarkable linearity of most of the data seriously raises the question of the correctness of the theoretical predictions. We state now the points that have occurred to us which might modify the previous picture.

1. The second-order perturbation theory does not apply, even at modest impurity concentrations.

2. The correct superconducting wave functions that should be used in the calculation do not have the spin-coherence properties attributed to those proposed by the B.C.S. theory.

3. The indirect spin–spin interaction via the conduction electrons lifts the degeneracy between different spin levels sufficiently so that an additional electron–electron interaction can arise which has an immediate effect on the energy gap. Such a situation might explain the greater-than-usual enhancement of the S.C. transition temperature of titanium with magnetic transition-element impurities of certain concentrations.

4. Crystal field splitting of the ionic spin levels, if it happens to have values of order $J^2\xi/E_f$, where J is the exchange integral between conduction electrons and f-shell electrons, and E_f is the Fermi energy, could similarly have a substantial effect on the energy gap. This would be in the nature of a coincidence.

5. As already indicated, there might be modifications of the electron wave functions due to the locally strong dipole fields of the ions, outside the range of validity of the usual perturbation theory.

Acknowledgement—We would like to thank Mr. L. D. LONGINOTTI for careful preparation of the many alloys used in this work.

* We wish to repeat that this method measures *volume* susceptibilities and will discriminate between a fully superconducting sample and superconducting impurities.

REFERENCES

1. MATTHIAS B. T., SUHL H. and CORENZWIT E., *Phys. Rev. Letters* **1**, 92 (1958).
2. SUHL H. and MATTHIAS B. T., *Phys. Rev.* **114**, 977 (1959).
3. MATTHIAS B. T., SUHL H. and CORENZWIT E., *Phys. Rev.* **115**, 1597 (1959).
4. MATTHIAS B. T. and HULM J. K., *Phys. Rev.* **87**, 799 (1952).
5. SUHL H. and MATTHIAS B. T., *Phys. Rev. Letters* **2**, 5 (1959).
6. MATTHIAS B. T., SUHL H. and CORENZWIT E., *J. Phys. Chem. Solids*, to be published.

VOL.
13
1960

LETTERS TO THE EDITOR

The limiting low-temperature behavior of the heat capacity of graphite*

(*Received* 16 *October* 1959)

DURING the past few years there has been wide experimental and theoretical interest in the heat capacity of layer-type crystals. Particular attention has been directed to graphite, and measurements on this substance have been made down to temperatures of about 1°K. Recently in this journal DE SORBO and NICHOLS[1] reported results for four different specimens of graphite for the temperature range 1–20°K. Interpretations of the results were given in accompanying papers by BOWMAN and KRUMHANSL[2] and by KOMATSU[3], the two approaches being essentially the same. The purposes of the present note are twofold: to show that there are inconsistencies in the analyses and interpretations of the results, and to propose an alternative approach.

Remarks on previous work

It seems to be well established[4, 5] that in the temperature region $15° < T < 50°$K the heat capacity of graphite shows at least approximately a T^2 temperature-dependence. Such a behavior can be accounted for theoretically (see, for example NEWELL[6]). Theory also predicts that in the limit of very low temperatures the heat capacity should become proportional to T^3. The new results obtained by DE SORBO and NICHOLS[1] are in accord with those published previously[4, 7] in showing an apparent limiting temperature-dependence lying between T^2 and T^3, the exact power depending upon the particular specimen. BOWMAN and KRUMHANSL[2] make a clear distinction between natural graphite (CNG) on the one hand and lampblack graphite (SA–25) on the other. The limiting heat capacity of the former is taken to be made up of linear electronic and cubic lattice terms. An additional contribution due to stacking faults is invoked to account for the larger heat capacity of lampblack graphite. KOMATSU[3] also introduces the concept of stacking faults to explain the apparent dependence of the heat capacity of graphite upon particle size.

DE SORBO and NICHOLS[1] describe their results for CNG in the temperature range 1·4–4·2°K by

$$C = 0{\cdot}6\times10^{-5}T + 4{\cdot}77\times10^{-6}T^3 \text{ cal/g. atom.deg} \qquad (1)$$

and remark that the coefficient of the linear term (γ) agrees well with that found previously by KEESOM and PEARLMAN[7] for a pile graphite. It is important to note, however, that the relation $C = \gamma T + aT^3$ is taken to hold over quite different temperature ranges for the two sets of experimental results. KEESOM and PEARLMAN's results depart strongly from the relation for $T > 2°$K, and their coefficient of the term in T^3 is about 60 per cent larger. A close examination shows that equation (1) does not represent DESORBO and NICHOLS's observed heat capacities within anything like their estimated accuracy (0·5 per cent between 1·5 and 4·2°K). There are systematic deviations of as much as 6 per cent, and they are in such directions as to indicate that the true graph of C/T against T^2 for CNG is a continuous curve with an intercept on the axis $T^2 = 0$ appreciably less than $0{\cdot}6\times10^{-5}$ cal/g. atom.deg^2. The claimed agreement of the electronic heat capacities is therefore illusory. In his interpretation of the heat-capacity results, KOMATSU has been obliged to use a value of γ of $0{\cdot}22\times10^{-5}$ cal/g.atom.deg^2, which is close to the value $0{\cdot}3\times10^{-5}$ cal/g.atom.deg^2 derived from a comprehensive theoretical calculation by MCCLURE[8]. Hence, the weight of evidence is for γ to be appreciably smaller than $0{\cdot}6\times10^{-5}$ cal/g.atom.deg^2.

The most striking feature of the experimental results is the "excess" heat capacity of the smaller-particle graphites. The explanation proposed by BOWMAN and KRUMHANSL[2] and by KOMATSU[3] supposes that stacking faults increase the proportion of vibrational modes excited at low temperatures. The result is an excess heat capacity which varies as T^2 and which gives rise to a different value of Θ_0 (Debye characteristic temperature at 0°K) for each of the specimens. While this explanation appears to be capable of accounting for the observed heat capacities, its consequences in the limit as T approaches zero have not been examined. The important point here is that

* National Research Council No. 5517.

a T^2 contribution due to stacking faults must disappear and the total lattice heat capacity go to a T^3 dependence in the limit; this cannot be seen unambiguously in the graphs of C/T against T^2. It will be shown below by analyzing the results in a different way that the T^2 term remains and therefore that the previous interpretations cannot be correct.

Fig. 1 is a graph of C/T^2 against T for the experimental results tabulated by DE SORBO and NICHOLLS[1] and by KEESOM and PEARLMAN[7]. The curves representing the different specimens are remarkably similar and asymptotically approach a consistent limiting heat capacity of the form

$$C_v = aT^3 + bT^2,$$

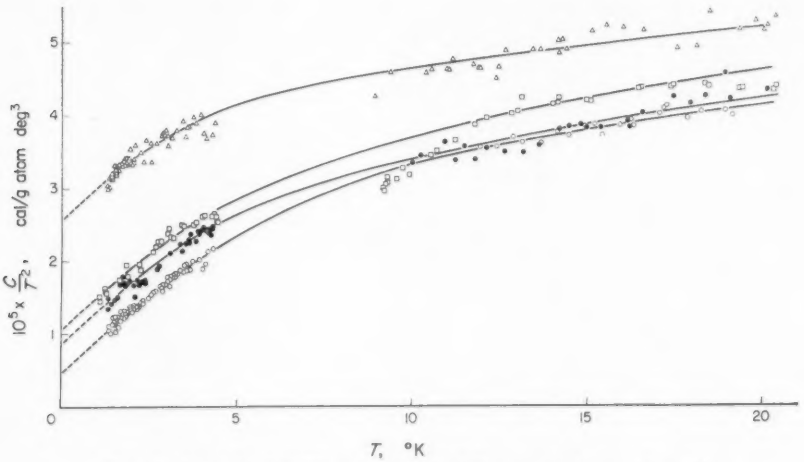


FIG. 1. A graph of C/T^2 against T for different specimens of graphite.

△—lampblack graphite (SA–25)
●—pile graphite (H–CS–II)
○—natural graphite (CNG)
} DE SORBO and NICHOLLS[1]

□—pile graphite. KEESOM and PEARLMAN[7]

A new interpretation

It is instructive to analyse the data in terms of a graph of C/T^2 against T. If the electronic contribution to the heat capacity is appreciable, it will be emphasized in a dramatic way by a sudden upward rise of the graph as T approaches zero. If, however, the electronic contribution is not appreciable, the limiting behaviour of the T^2 term will be clearly brought out. If the explanation of BOWMAN and KRUMHANSL[2] and KOMATSU[3] for the excess heat capacity of the smaller-particle graphites is correct, C/T^2 should go to zero at $T = 0$°K. Of course, a combination of contributions due to an electronic term and to stacking faults could complicate the limiting behaviour, but this should be immediately apparent from a lack of regularity in the results for different graphite specimens.

in which the coefficient a is essentially the same for all of the specimens. The contribution in T^2 clearly remains in the limit and must therefore have a different origin from that suggested by BOWMAN and KRUMHANSL[2] and by KOMATSU[3]. An obvious alternative is that the contribution arises from a boundary effect such as has been discussed by MONTROLL[9]. For the graphitized lampblack specimen for which BOWMAN and KRUMHANSL[2] estimate the mean particle size to be about 100 Å, we obtain $b \sim 1{\cdot}0 \times 10^{-5}$ cal/g. atom. deg^3; this is to be compared with $b \sim 2{\cdot}5 \times 10^{-5}$ cal/g.atom.deg^3 taken from Fig. 1. The agreement is satisfactory, since the value of b calculated on the basis of MONTROLL'S theory may be in error by as much as a factor of two.[10] Essentially the same calculation was done by BOWMAN and KRUMHANSL (Ref. 2, p. 378), but

they discounted the particle-size effect because "the 'surface' modes would not *predominate* over the volume modes unless T were below about 2°K".

While the comparison of absolute values is satisfactory for the graphitized lampblack, there are, at first sight, difficulties with the other specimens. The average dimension of the pile-graphite crystallites was estimated to be 200–300 Å, and so we should expect b for these specimens to be at most about one-quarter of that for graphitized lampblack. Fig. 1 indicates that it is more nearly one-half. Also, the heat capacity of natural graphite (CNG) appears to show an appreciable particle-size effect, and this cannot be reconciled with an estimated average crystal size of 60–100μ. Now, in the immediately foregoing discussion the electronic contribution to the heat capacity has been totally neglected. It is true that no upward curvature of the graph of C/T^2 as T approaches zero can be seen in Fig. 1, but this does not rule out the possibility that there is a small electronic contribution. In fact, calculation shows that results well below 1°K would need to be obtained before a contribution corresponding to $\gamma = 0{\cdot}3\times10^{-5}$ cal/g.atom.deg^2 could be seen unambiguously. We have therefore investigated whether or not the apparent discrepancies in the particle-size effect can be removed by introducing a small electronic term.*

Fig. 2 is a graph of $(C-\gamma T)/T^2$ against T for all of the results, where we have assumed the value $\gamma = 0{\cdot}3\times10^{-5}$ cal/g.atom.deg^2 obtained from theory.[8] This small modification is enough to remove the inconsistencies in interpreting the T^2 contribution to the heat capacity as a boundary effect. In particular, no boundary effect is now

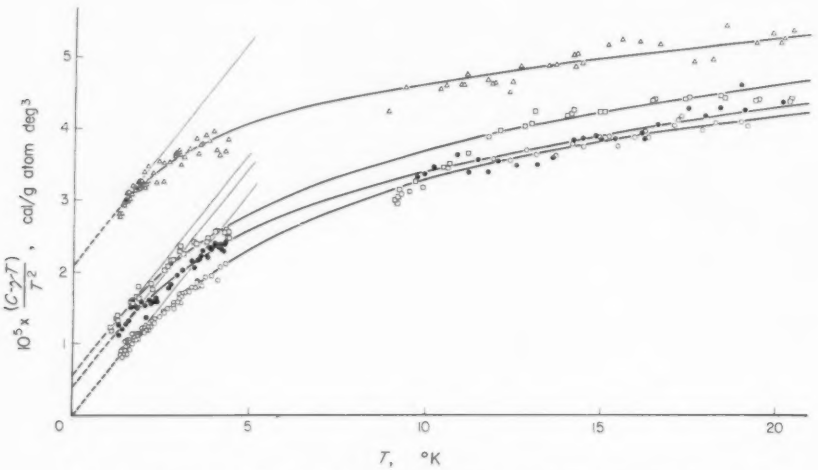


FIG. 2. A graph of $(C-\gamma T)/T^2$ against T for different specimens of graphite with $\gamma = 0{\cdot}3\times10^{-5}$ cal/g. atom deg^2. Symbols as in Fig. 1.

observed in the heat capacity of the natural graphite, and the coefficients of the T^2 terms for the small-particle graphites are approximately proportional to the estimated specific surfaces of the specimens. Also, the identity of the coefficients of the T^3 terms for the different graphites is unaltered. The lighter lines in Fig. 2, which represent the limiting slopes of the curves drawn through the experimental points, give $a = 6{\cdot}2_7 \pm 0{\cdot}06\times10^{-6}$ cal/g.atom.deg^4, which corresponds to $\Theta_0 = 420\pm10$°K.

Summary

Previous interpretations of the low-temperature heat capacity of graphite have been critically examined and shown to contain inconsistencies.

* This was suggested to us by Professor KRUMHANSL.

By a different analysis of the available experimental results, it has been demonstrated that the *limiting* low-temperature heat capacity of graphite can be represented by

$$C = 0{\cdot}3\times10^{-5}T+6{\cdot}2_7(\pm0{\cdot}06)\times10^{-6}T^3+bT^2$$

cal/g.atom.deg,

where bT^2 represents the effect of particle size on the heat capacity. It is therefore evident that, contrary to previous work, no special effects need to be invoked to describe the low-frequency part of the lattice-frequency spectrum of graphite. The "bulk" properties of different graphite specimens are identical within experimental uncertainty.

The existence of the effect of particle size on the heat capacity complicates the usual type of analysis in terms of C/T against T^2 for the electronic contribution to the heat capacity. While it turns out that an accurate value of γ cannot be obtained directly from the existing experimental results for graphite, the results are consistent with $\gamma = 0{\cdot}3\times10^{-5}$ cal/g.atom.deg^2 which has been obtained from theory.

Acknowledgement—We should like to thank Dr. T. H. K. BARRON for useful comments and to acknowledge a helpful discussion with Professor J. A. KRUMHANSL.

Division of Pure Chemistry P. FLUBACHER*
National Research Council A. J. LEADBETTER*
Ottawa, Canada J. A. MORRISON

References

1. DE SORBO W. and NICHOLS G. E., *J. Phys. Chem. Solids* **6**, 352 (1958).
2. BOWMAN J. C. and KRUMHANSL J. A., *J. Phys. Chem. Solids* **6**, 367 (1958).
3. KOMATSU K., *J. Phys. Chem. Solids* **6**, 380 (1958).
4. BERGENLID U., HILL R. W., WEBB F. J. and WILKS J., *Phil. Mag.* **45**, 851 (1954).
5. DE SORBO W. and TYLER W. W. *J. Chem. Phys.* **21**, 1660 (1953).
6. NEWELL G. F., *J. Chem. Phys.* **23**, 2431 (1955).
7. KEESOM P. H. and PEARLMAN N., *Phys. Rev.* **99**, 1119 (1955).
8. MCCLURE J. W., *Phys. Rev.* **108**, 612 (1957).
9. MONTROLL E. W., *J. Chem. Phys.* **18**, 183 (1950).
10. STRATTON R., *Phil. Mag.* **2**, 702 (1957).

* National Research Council Postdoctorate Research Fellow.

Absorption edge in β-silicon carbide

(*Received* 18 *September* 1959;
revised 30 *October* 1959)

THE optical absorption of a β-SiC crystal has been measured at 20°C in the range 2·3–2·9 eV. The single crystal used was from a lot grown by the KENDALL[1] method. X-ray powder pattern analysis of this lot showed that it consisted entirely of the cubic modification. The sample was sectioned and polished to give a plate approximately 0·3 mm thick and 2 mm in dia. Transmission spectra of this specimen were taken at three successive thicknesses (0·257, 0·174 and 0·113 mm) produced by polishing. The sample was polished by standard methods, but the faces were parallel only to within ±3 per cent. Thicknesses were measured with a calibrated microscope.

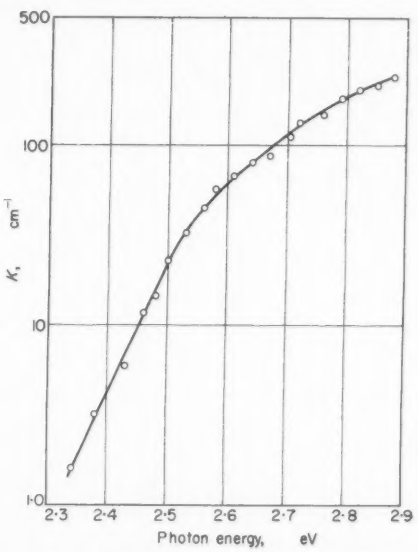


FIG. 1. Absorption spectrum of β-SiC.

Spectra were obtained by means of a Leiss single-prism monochromator, an RCA 7265 high-gain photomultiplier tube and a General Radio electrometer. The resolution was about 0·01 eV in the region studied. The region from 2·3 to 2·5 eV was not studied for the largest sample thickness.

From the relation $T = F(R)\exp(-Kt)$, where T is the transmissivity, $F(R)$ a function of the reflectivity R, K the absorption coefficient and t the

sample thickness, it is seen that a plot of ln T vs. t for a given incident photon energy gives a straight line of slope K, the absorption coefficient at that photon energy. Values of T were read from the transmission curves at intervals of 5 mμ. The slopes of the straight lines obtained in this manner determined the absorption curve in Fig. 1.

In order to estimate the value of the energy gap E_G, an extrapolation of the absorption curve to $K = 0$ was made. KOBAYASI's[2] recent calculation

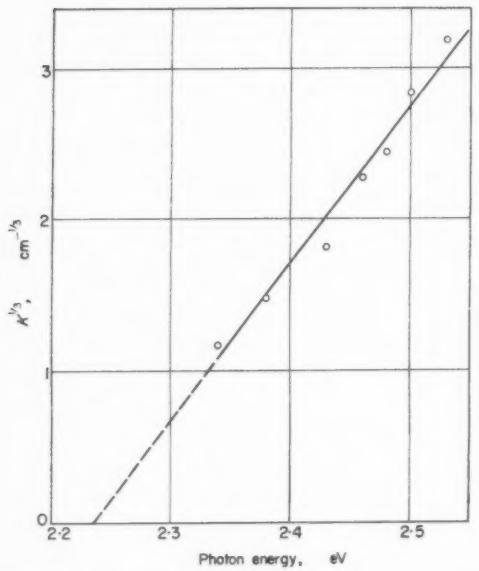


FIG. 2. Cube root of the absorption coefficient vs. photon energy.

of the energy-band structure of β-SiC suggests the presence of indirect transitions.[3] If such transitions do occur, then extrapolation of a plot of either $K^{1/2}$ or $K^{1/3}$ vs. $h\nu$ offers a method of estimating E_G. Since the plot of $K^{1/2}$ vs. $h\nu$ was non-linear, an extrapolation of the linear plot of $K^{1/3}$ vs. $h\nu$ was used. This is shown in Fig. 2, and it is seen that the intercept on the $h\nu$-axis is 2·24 eV. A value of 2·2 eV is therefore suggested for the optical band gap. The estimated experimental error is $\pm$ 0·1 eV.

This value may be compared with the result given by POHL[4], based on electrical measurements, that $E_G = 1{\cdot}90 \pm 0{\cdot}10$ eV. LELY and KRÖGER[5] report $E_G \sim 2{\cdot}8$ eV, based on a room-temperature transmission curve which shows a 440 mμ cut-off. Although PHILIPP[6] gives no value for E_G, he reports $K = 100$ cm^{-1} for $h\nu = 2{\cdot}62$ eV, which agrees well with the value of 2·68 eV from Fig. 1. The calculated value[2] of the long-wave limit of indirect transitions is 2·2 eV.

The resolution was insufficient for the detection of possible fine structure[7] in the absorption edge due to indirect transitions. There is no break in the curve suggestive of the onset of direct transitions, which KOBAYASI[2] calculates will take place at a photon energy of 6·3 eV. Further work at higher resolution and larger values of K will be valuable in connection with these questions.

Acknowledgement—The author would like to acknowledge several valuable discussions with Mr. R. ZALLEN and to thank Mr. R. C. ELLIS, Jr., for furnishing the single crystal of β-SiC.

Research Division RICHARD DALVEN
Raytheon Company,
Waltham 54, *Mass.*

References

1. KENDALL J. T., *J. Chem. Phys.* **21**, 821 (1953).
2. KOBAYASI S., *J. Phys. Soc. Japan* **13**, 261 (1958).*
3. HALL L. H., BARDEEN J. and BLATT F. J., *Phys. Rev.* **95**, 559 (1954).
4. POHL R. G., Program of the Conference on Silicon Carbide, Boston, 2–3 April, 1959.
5. LELY J. A. and KRÖGER F. A., *Semiconductors and Phosphors*, Proceedings of the International Colloquium, 1956, at Garmisch–Partenkirchen, p. 514. Vieweg, Braunschweig (1958).
6. PHILIPP H. R., *Phys. Rev.* **111**, 440 (1958).
7. MACFARLANE G. G., MCLEAN T. P., QUARRINGTON J. E. and ROBERTS V., *Phys. Rev.* **108**, 1377 (1957).

The temperature-dependence of electron spin-lattice relaxation times in ruby

(*Received* 12 *November* 1959; *revised* 9 *December* 1959)

THE spin-lattice relaxation time, T_1, for the Cr^{3+} ion has been measured as a function of temperature in single crystals of ruby at 9·3 kMc/s. The energy

* The author would like to thank the referee for bringing this paper to his attention.

states for the Cr^{3+} ion in this case can be described by the spin Hamiltonian[1]

$$\mathcal{H} = D\,S_z^2+\beta[g_\parallel H_z+g_\perp(H_x+H_y)-1/3\;S(S+1)]$$

with $g_\parallel = 1{\cdot}982$, $g_\perp = 1{\cdot}979$ and $D = 0{\cdot}1912$ cm^{-1}.

Values of T_1 were obtained by observing the signal saturation as the microwave power incident on the cavity was increased.[2] Measurements were made on the 2→3 transition with the magnetic field at an angle of 50° to the trigonal axis. The energy levels are numbered from lowest to highest energy values and are shown in Fig. 1 for this orientation. Measurements were made in the temperature region 77–300°K with crystals containing 0·65 and 0·055 per cent Cr_2O_3. The widths between inflection points of the unsaturated absorption curves, for the two concentrations, were 50 and 20 G respectively. The detector crystal was maintained at a fixed operating point by means of a pair of calibrated attenuators. Samples were cooled by conduction through a sapphire rod in contact with a constant-temperature bath, the temperature of the sample being measured by a thermocouple.

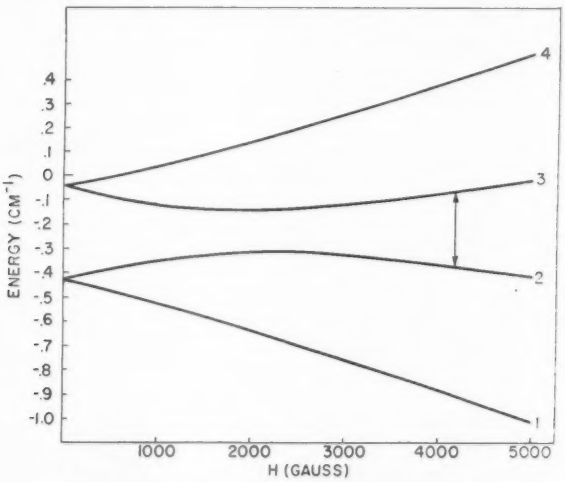


FIG. 1. Energy levels of the Cr^{3+} ion in Al_2O_3 for $\theta = 50°$. The arrow indicates the transition observed.

VAN VLECK[3] has considered the spin-lattice relaxation of electronic states by means of spin-orbit coupling and lattice vibrations. His results indicate that at very low temperatures the energy exchange involves a single lattice phonon with a frequency equal to the electronic frequency. The predicted temperature-dependence of the relaxation time for this process is $1/T$. As the temperature is raised, more high-frequency lattice oscillators are excited, and a double phonon, or Raman-like interaction, dominates as the relaxation mechanism. This is expected to give rise to a stronger temperature-dependence for T_1. The present work was undertaken to determine at what temperature the Raman process begins to dominate the single-phonon one.

Previous measurements[4] on the more heavily doped sample at 4·2 and 77°K indicated a $1/T$-dependence in this region. Other measurements[5] on ruby in the liquid-nitrogen region also showed a $1/T$-dependence for T_1. The measurements reported in the present work are shown as the points plotted in Fig. 2, along with the theoretical temperature variation. The previously determined $1/T$ dependence was fitted to the value of T_1 at 77°K. The slope of the steeper curve, corresponding to a dominant Raman process, was obtained by numerical integration of VAN VLECK's[3] temperature-dependent factor for a Debye temperature of 980°K. The ordinate of the curve was fitted to the measured value of T_1 at 240°K. Both crystals

show reasonable agreement with the expected Raman temperature-dependence. The results indicate that for ruby the Raman process is dominant at temperatures greater than approximately 150°K,

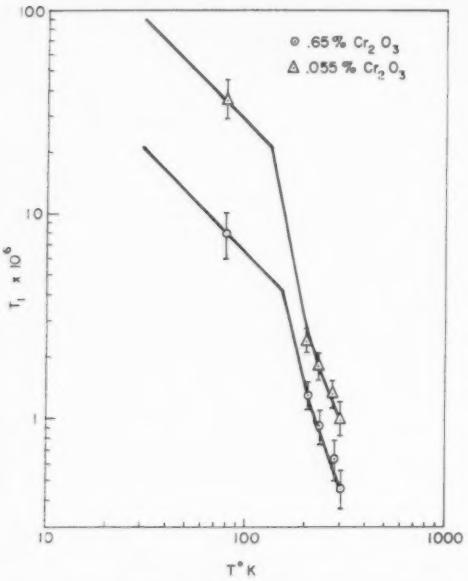


FIG. 2. Experimental and theoretical temperature-dependence of electron spin-lattice relaxation times in ruby.

the single-phonon process being dominant below this temperature. The fact that the Raman process is the dominant one over only a portion of the region between 77 and 300°K explains the relatively slow dependence for T_1 based on measurements made at these points, namely

$$T_1 \propto 1/T^{2.5 \pm 0.5}.$$

If the Raman process were dominant over the entire region, one would expect

$$T_1 \propto 1/T^{5.5}.$$

The results of this work are in disagreement with measurements made by PASHININ and PROKHOROV[6]. They published values of T_1 for ruby at 77 and 300°K which indicate $T_1 \propto 1/T^{6.3}$. Their measurements were made on the $2 \rightarrow 4$ transition for $\theta = 0°$. Although it is known that T_1 is orientation-dependent, there is evidence that the temperature-dependence is only slightly if at all dependent upon orientation.[4] Measurements[4] made on our more heavily doped crystal for the same orientation, transition and temperature as those PASHININ and PROKHOROV used are in agreement with the results reported here and in disagreement with their results.

It should be noted that the concentration-dependence of T_1 is not explained by VAN VLECK's theory. Since there is a concentration-dependence even at temperatures approximately one-third of the Debye temperature, it is not reasonable to associate the dependence with a phonon "bottleneck".[7] It seems that the energy exchange between the electron and the lattice must itself depend upon the concentration of the paramagnetic impurity.

R.C.A. Laboratories R. E. MICHEL
Princeton, N.J.

References

1. MANENKOV A. A. and PROKHOROV A. M., *J. Exp. Theor. Phys.* (*U.S.S.R.*) **28**, 762 (1955); *Sov. Phys. JETP* **1**, 611 (1955).
2. BLOEMBERGEN N., PURCELL E. M. and POUND R. V., *Phys. Rev.* **73**, 679 (1958).
3. VAN VLECK J. H., *Phys. Rev.* **57**, 426 (1940).
4. LEWIS H. R. and MICHEL R. E., *Bull. Amer. Phys. Soc.* **3**, 179 (1958).
5. Armour Research Foundation Contract with Evans Signal Laboratory No. 36-039-SC-73032, Final Report No. 1 (January 1, 1958).
6. PASHININ P. P. and PROKHOROV A. M., *J. Exp. Theor. Phys.* (*U.S.S.R.*) **34**, 777 (1958); *Sov. Phys. JETP* **7**, 535 (1958).
7. GIORDMAINE J. A., ALSOP L. E., NASH F. R. and TOWNES C. H., *Phys. Rev.* **109**, 302 (1958).

Neutron-diffraction Study of Ti_2O_3*

(*Received* 1 *October* 1959; *revised* 9 *November* 1959)

IN recent years an extensive study has been carried out on the magnetic properties of rhombohedral sesquioxides such as $\alpha-Fe_2O_3$ and Cr_2O_3 and the antiferromagnetic ordering of these compounds was determined by neutron-diffraction analysis. Among several other isomorphous compounds of

* This work was carried out in part at Brookhaven National Laboratory under the auspices of the U.S. Atomic Energy Commission and the National Security Agency.

VOL. 13 1960

the corundum-type structure, Ti_2O_3 and V_2O_3 were reported to be antiferromagnetic. The magnetic properties of Ti_2O_3 were investigated thoroughly by PEARSON[1], and the general increase of the susceptibility with the temperature was interpreted as an indication of antiferromagnetic ordering at room temperature. The existence of a phase transition near 200°C was indicated by a sudden increase in the susceptibility and by a rapid change of lattice constants in the same temperature range. PEARSON pointed out that the change of susceptibility near 200°C is similar to that found in $\alpha-Fe_2O_3$ at −20°C, where a change in the direction of the magnetic axis takes place. MORIN[2], on the other hand, considered this phase transition as the Néel point because of the large change of the electrical conductivity observed.

A neutron-diffraction study of Ti_2O_3 was initiated in order to give positive evidence of the antiferromagnetic ordering. Approximately 30 g of Ti_2O_3 powder were prepared by Dr. W. D. JOHNSTON, of the Westinghouse Research Laboratories, by reduction of TiO_2 with titanium metal. The reducing power of Ti^{3+} in the compound is 99·1 per cent of the theoretical value, and the lattice constants are in good agreement with those given by PEARSON. The general tendency of the temperature-dependence of the susceptibility agrees with the previous report.[1]

Neutron-diffraction data were obtained at 295 and 4·2°K. No noticeable difference was observed between the two sets of data. No peaks were observed at (111) and (100), which should be entirely magnetic in origin. The observed intensities can be explained quite satisfactorily by assuming positional parameters of $u = 0{\cdot}345$ for titanium and $x = 0{\cdot}565$ for oxygen. These values were obtained by the single-crystal X-ray measurements.[3] It can be concluded that the magnetic intensities, if they exist, must be quite weak.

The maximum magnetic moment for an antiferromagnetic ordering can be estimated from the observable limits of our powder intensities on (111) (110) and (110). We consider three probable models for the antiferromagnetic ordering, proposed by SHULL *et al.*[4] for $\alpha-Fe_2O_3$, which can be obtained by changing the sequence of the sign of the Ti^{3+} moment along the rhombohedral axis as (++−−), (+−+−) and (+−−+). From our experimental accuracy it can be concluded that the moment of Ti^{3+} must be smaller than $0{\cdot}5\mu_B$ for any spin direction, if we consider the three models with a single spin axis.

Preliminary measurements were also carried out with single crystals. The crystals were rotated around the $[1\bar{1}0]$ axis, and the intensities at (111), (100) and (110) were compared with the intensities at (222), (200) and (220) respectively. Because of a probable extinction effect and the contamination of half wavelength, the existence of weak magnetic intensities could not be ruled out at present. However, the maximum value for the possible moment can be set at $0{\cdot}3\ \mu_B$.

Since the expected value of the magnetic moment of Ti^{3+} is $1\ \mu_B$, our experimental results indicate either that Ti_2O_3 is not antiferromagnetically ordered or that, if it is ordered, the moment must be reduced considerably. Very recently, GOODENOUGH[5] has proposed a theory based upon direct cation–cation interactions which explains the transition near 200°C as a result of an ordering of electrons into a covalent-type bond. This theory predicts that the localized atomic moment of Ti_2O_3 is extremely small and that long-range antiferromagnetic order may exist only at very low temperatures. Our results are compatible with this bonding theory of GOODENOUGH, which is also supported by the recent susceptibility measurements on Ti_2O_3 single crystals by CARR and FONER[6]. It may be interesting to note that a neutron-diffraction study of V_2O_3, recently carried out by PAOLETTI and PICKART[7], did not give any positive evidence of antiferromagnetism.

Acknowledgements—We are very grateful to Dr. W. D. JOHNSTON for helpful discussions on the crystal chemistry and for supplying the powder sample, and to Dr. R. R. HEIKES, for discussion on the magnetic properties and for providing the magnetic data. We are also indebted to Dr. J. B. GOODENOUGH for the communication prior to publication of his illuminating results.

Westinghouse Research Laboratories Pittsburgh, Pa., G. SHIRANE

Naval Ordnance Laboratory, White Oak, Md., and Brookhaven National Laboratory, Upton, N.Y. S. J. PICKART

Laboratory for Insulation Research, Massachusetts Institute of Technology, Cambridge, Mass. R. NEWNHAM

References

1. PEARSON A. D., *J. Phys. Chem. Solids* **5**, 316 (1958).
2. MORIN F. J. *Phys. Rev. Letters* **3**, 34 (1959).
3. NEWNHAM R. Unpublished.
4. SHULL C. G., STRAUSER W. A. and WOLLAN E. O. *Phys. Rev.* **83**, 333 (1951).
5. GOODENOUGH J. B. Private communication. Abstract of the Conference on Magnetism and Magnetic Materials, Detroit (Nov. 1959).
6. CARR P. H. and FONER S. Abstract of the same Conference.
7. PAOLETTI A. and PICKART S. J., *J. Chem. Phys.* (to be published).

Jonctions de germanium *p–n* plastiquement deformées

(*Received* 4 *September* 1959)

LES jonctions de germanium *p–n* en polarisation inverse présentent souvent le phénomène dit "claquage doux" (*soft breakdown*); ce phénomène peut, dans certains cas particuliers, être attribué à des effets de surface; dans les autres cas il ne semble pas qu'une explication satisfaisante ait été proposée. Les expériences décrites ci-dessous montrent que les dislocations peuvent, au moins dans certains cas, être responsables de ce phénomène.

Dans un cristal de germanium *p–n* obtenu par tirage on a découpé deux bâtonnets identiques dans une région du cristal pratiquement exempte de dislocation; la caractéristique inverse de la première jonction a été mesurée à trois températures fixes 78, 195 et 273°K (voir Fig. 1); le courant inverse bien saturé est en bon accord avec les théories de la jonction *p–n*;[1–3] le claquage est brusque (*abrupt breakdown*) et en bon accord avec les résultats connus de l'avalanche par multiplication;[4, 5] on a vérifié que le photo-courant est à peu près multiplié dans le même rapport. Le deuxième bâtonnet contenant la jonction a été déformé plastiquement à 600°C par flexion entre deux formes de quartz d'un rayon de courbure de 5 cm; le nombre de dislocations introduites au niveau de la jonction est de l'ordre de 10.000; les caractéristiques inverses de cette deuxième jonction, mesurées aux mêmes températures (78, 195 et 273°K) sont portées sur la Fig. 1; on peut faire les constatations suivantes :

Il ressort de ces expériences que des dislocations au niveau d'une jonction *p–n* introduisent un courant excédentaire, croissant très rapidement avec le champ électrique de la zone chargée et à peu près indépendant de la température; il est naturel de penser que ce courant est dû à une émission de champ interne; deux mécanismes semblent possibles : les charges électriques dues à des électrons ou des trous piégés au voisinage des

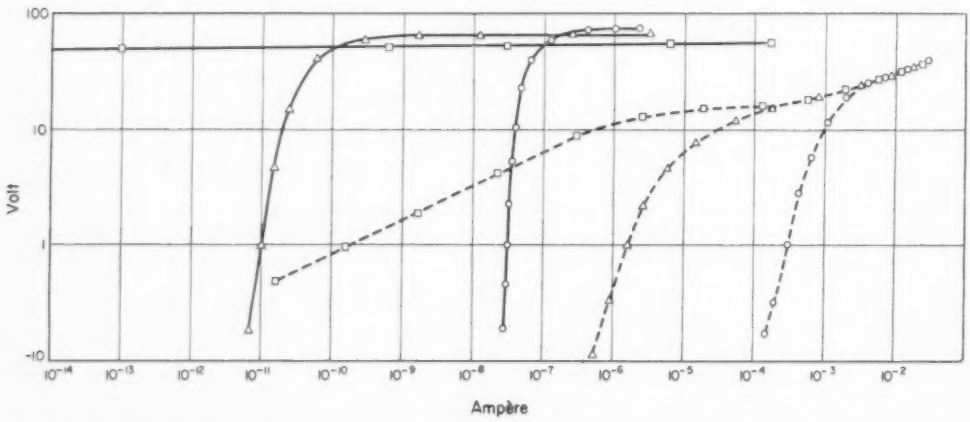


FIG. 1. Caracteristiques inverses de deux jonctions de germanium à différentes températures: □78°K, △195°K, ⊙273°K; les courbes en trait plein correspondent à une jonction exempte de dislocations, les courbes en traits interrompus à une junction plastiquement déformée et contenant environ 10.000 dislocations.

dislocations peuvent créer des zones où le champ électrique local excède nettement le champ électrique moyen de sorte que se trouvent localement remplies des conditions favorables à un effet tunnel à travers la bande interdite (effet Zener). Il existe une autre possibilité : le champ de contraintes de la dislocation peut entraîner localement une diminution de la bande interdite[6] qui rend possible un effet Zener avec la valeur moyenne du champ électrique qui règne (environ 10^5 V/cm).

Des expériences plus complètes sont en cours pour préciser ce phénomène et décider quel est le mécanisme effectivement responsable; ces expériences feront l'objet d'un prochain article.

Centre National d'Etudes des Télécommunications
Issy-les-Moulineaux (Seine)
France.

M. BERNARD
B. LEDUC

VOL. 13 1960

Bibliographie

1. SHOCKLEY W., *Bell Syst. Tech. J.* **28**, 435 (1949).
2. BERNARD M., *Semiconducteurs et Phosphores* p. 344 (1956); *J. Electronics* **2**, 579 (1957).
3. SAH C. T., NOYCE R. N. et SHOCKLEY W., *Proc. Inst. Radio Engrs.* **45**, 1228 (1957).
4. MCKAY K. G. et MCAFFEE K. B., *Phys. Rev.* **91**, 1079 (1953).
5. MILLER S. L. *Phys. Rev.* **99**, 1234 (1955).
6. KULIN S. A. et KURTZ A. D., *Acta. Met.* **2**, 354 (1954).

CONTENTS LISTS OF VOLUME 9, NOS. 1–4

Tôme 9 Nº 1

page

A. A. MARADUDIN:
Dislocations vis et théorie élastique discrète 1

G. A. ALERS, J. R. NEIGHBOURS et H. SATO:
Influence de la direction d'aimantation dans le nickel aux champs éléves sur la vitesse du son et son amortissement 21

G. L. PEARSON et W. L. FELDMANN:
Techniques des poudres pour tracer les domaines ferro-électriques 28

D. G. THOMAS:
Diffusion et précipitation de l'indium dans l'oxyde de zinc. 31

M. J. STEPHEN et A. B. LIDARD:
L'effet Faraday dans les semi-conducteurs 43

H. E. FARNSWORTH et J. TUUL:
Détérmination par la diffraction des électrons de basse énergie de l'influence de la densité des défauts de réseau sur la chimisorption et la formation d'oxyde sur une face cristalline (100) propre de nickel 48

H. E. FARNSWORTH, R. E. SCHLIER et J. TUUL:
Détection par la diffraction des électrons de basse énergie de l'hydrogène absorbé et de l'hydroxyde de nickel formé sur la surface (100) d'un cristal de nickel à 25°C . . 57

F. A. TRUMBORE, C. R. ISENBERG et E. M. PORBANSKY:
Sur la variation du coefficient de distribution avec la température. Les solubilités à l'état solide de l'étain dans le silicium et le germanium 60

W. KÄNZIG et T. O. WOODRUFF:
La structure électronique d'un centre H 70

Lettres aux Editeur:
E. CORENZWIT: Supraconductibilité de Nb_3Al 93

M. C. STEELE: Rupture électrique à champ faible dans le phosphure d'indium de type *n* 93

Tôme 9 Nº 2

T. P. HOAR et N. F. MOTT:
Un mécanisme de formation de films d'oxyde anodiques poreux sur l'aluminium . . 97

R. E. SCHMUNK et C. S. SMITH:
Dérivées par rapport à la pression des constantes élastiques de l'aluminium et du magnésium 100

H. C. NASH et C. S. SMITH:
Constantes élastiques du lithium monocristallin 113

R. T. BATE, R. K. WILLARDSON et A. C. BEER:
Magnétorésistance transverse et effet Hall dan InSb de type *N* 119

H. EHRENREICH:
Transport des électrons dans InSb intrinsèque 129

N. T. MELAMED:
Transfert d'énergie dans les composés phosphorescents ZnS:Cu:In 149

VOL. 13 1960

page

T. Riste, K. Blinowski et J. Janik:
Diffusion des neutrons par les fluctuations de spin dans la magnétite 153

G. H. Jonker:
Analyse des propriétés semi-conductrices du ferrite de cobalt 165

A. Wold et R. J. Arnott:
Préparation et propriétés cristallographiques des systèmes $LaMn_{1-x}Mn_xO_{3+\lambda}$ et $LaMn_{1-x}Ni_xO_{3+\lambda}$ 176

Lettres aux Editeur:

I. Broser et R. Broser-Warminsky: Réponse à la remarque de M. Balkanski sur l'article "Sur le mécanisme du déplacement de l'énergie dans les cristaux de sulfure de cadmium" 181

E. Brian Smith: Coexistence de phases et hystérésis 182

A. Giansoldati, J. O. Linde et G. Borelius: Les tranformations antiferromagnétiques de l'alliage AuMn 183

Tôme 9 Nos 3/4

H. A. Klasens:
L'influence de la température sur la fluorescence des photoconducteurs 185

S. W. Barber et B. Martin:
Une fonction tenant compte de la dispersion des vitesses pour la chaleur spécifique des solides isotropes 198

H. J. Hrostowski et R. H. Kaiser:
La solubilité de l'oxygène dans le silicium 214

F. Moser, N. R. Nail et F. Urbach:
Etude par absorption optique de la photolyse de volume de grands cristaux de chlorure d'argent 217

S. Geller et M. A. Gilleo:
L'effet des corrections de dispersion sur le raffinement de la structure du grenat fer-yttrium 235

R. S. Knox:
Etats excitoniques dans les cristaux de gaz rares 238

R. S. Knox:
Etats excitoniques dans l'argon cristallin 265

W. F. Love:
Effet Hall dans le cuivre et Cu_3Au à basse température 281

H. Meyer et P. L. Smith:
Les chaleurs spécifiques de sept éthylsulfates de terres rares aux basses températures . 285

H. Meyer:
Propriétés magnétiques et thermiques de l'éthylsulfate de praséodyme audessous de 1°K. 296

R. J. Weiss et A. S. Marotta:
Variation avec le spin de la résistivité des métaux magnétiques 302

G. J. Wolga et M. W. P. Strandberg:
Interactions hyperfine dans les spectres de résonance magnétique de centres F dans LiF et NaF 309

page

S. Zwerdling, B. Lax, K. J. Button et Laura M. Roth:
Magneto-absorption oscillatoire dans l'antimonure de gallium 320

W. D. Lawson, S. Nielsen, E. H. Putley et A. S. Young:
Préparation et propriétés de HgTe et des cristaux mixtes de HgTe—CdTe 325

S. Wiederhorn et H. G. Drickamer:
L'influence de la pression sur le spectre le quelques cristaux aromatiques condensés dans l'ultra-violet proche 330

Lettres aux Editeur:
Keiji Maeda: Courbes en coordonnées de configuration unidimensionnelles pour les centres luminescents 335

M. J. Arvin, R. F. Tipsord: Coefficient de Hall du siliciure de molybdène 336

Revue de Livre:
L. R. G. Treloar: "La physique de l'élasticité du caoutchouc" 338

Том 9 Но. 1

А. А. Марадудин:
Винтовые дислокации и дискретная теория упругости 1

Г. А. Алерс, Дж. Нейборс и Г. Сато:
Зависимость скорости и поглощения звука от направления намагниченности никкеля в сильных полях 21

Г. Л. Пирсон и В. Л. Фельдман:
Порошковый метод для обнаружения ферроэлектрической доменной структуры . 28

Д. Томас:
Диффузия и выпадение индия из окиси цинка 31

М. Стивенс и А. В. Лидард:
Эффект Фарадея в полупроводниках 43

Г. Е. Фарнсворт и И. Туул:
Влияние плотности дефектов решетки на хемисорбцию и образование окислов на чистой кристаллической грани (100) никкеля, определенное по диффракции электронов малой энергии 48

Г. Фарнсворт, Р. Шлир и И. Туул:
Анализ с помощью диффракции электронов малой энергии адсобированного водорода и гидрада окиси никкеля на новерхности (100) кристалла никкеля при 25°С 57

Ф. А. Трумбор, К. Айзенберг и Е. М. Порбанский:
О температурной зависимости коэффициента распределения. Растворимости олова в твердых растворах в германии и кремнии 60

В. Кенциг и Т. О. Вудрафф:
Электронная структура Н-центра 70

Письма в Редакцию:

Е. Коренцвит: Сверхпроводимость Nb_3Al 93

М. Стиил: Низковольтный электрический пробой в *n*-фосфиде индия. . . . 93

VOL. 13 1960

Том 9 Но. 2

page

Т. П. Хор и Н. Ф. Мотт:
Механизм образования пористых анодных пленок окисла на алюминии . . 97

Р. Е. Шмунк и Ч. С. Смит:
Производные по давлению от упругих постоянных алюминия и магния . . 100

Г. Неш и Ч. Смит:
Упругие константы монокристалла лития 113

Р. Т. Бейт, Р. К. Виллардсон и А. Беер:
Поперечное магнитосопротивление и эффект Холла в InSb *n*-типа . . . 119

Г. Эренрейх:
Перенос электронов в InSb 129

Н. Т. Меламед:
Перенос энергии в фосфорах ZnS:Cu:In 149

Т. Ристе, К. Блиновский и Е. Яник:
Рассеяние нейтронов на сициоых флуктуациях в магнетите 153

Г. Джонкер:
Анализ полупроводниковых свойств феррита кобальта 165

А. Волд и Р. Арнотт:
Приготовление и кристаллографические свойства систем $LaMn_{1-x}Mn_xO_{3+\lambda}$ и $LaMn_{1-x}Ni_xO_{3+\lambda}$ 176

Письма в Редакцию:

И. Брозер, И. Брозер-Варминский: Ответ на замечание М. Балканского по поводу работы «О механизме миграции энергии в кристаллах сульфида калия» . 181

Е. Б. Смит: Сосуществование фаз и гистерезис 182

А. Джиансолдати, Дж. Линде, Г. Борелиус: Антиферромагнитные превращения в сплаве AuMn 183

Том 9 На. 3/4

Г. А. Класенс:
Температурная зависимость флуоресценции фотопроводников . . . 185

С. В. Барбер и Б. Мартин:
Теплоемкость изотропных твердых тел с учетом дисперсии скорости звука . . 198

Г. Хростовский и Р. Кайзер:
Растворимость кислорода в кремнии 214

Ф. Мозер, Н. Р. Нейл и Ф. Урбах:
Оптико-абсорбционное исследование объемного фотолиза в больших кристаллах хлорида серебра 217

С. Геллер и М. А. Гиллео:
Влияние дисперсионных поправок на уточнение структуры иттрий-железных гранатов 235

Р. С. Нокс:
Экситонные состояния в кристаллическом аргоне 238

page

Р. С. Нокс:
Экситонные состояния в кристаллах редких газов 265

В. Ф. Лав:
Холл-эффект в меди и в Cu_3Au при низких температурах 281

Х. Мейер и П. Л. Смит:
Теплоемкость семи этил-сульфатов редких земель при низких температурах . 285

Х. Мейер:
Тепловые и магнитные свойства этилсульфада празеодима при температурах ниже 1°К. 296

Р. Вейсс, А. С. Маротта:
Спиновая засисимость сопротивления магнитных металлов 302

Г. Волга, М. В. П. Страндберг:
Сверхтонкие взаимодействия в спектре парамагнитного резонанса F-центров в LiF и NaF 309

С. Цвердлинг, Б. Лакс, К. Баттон, Л. Рот:
Осциллирующее магнитное поглощение в антимониде галлия JA-1149 . . 320

В. Д. Лоусон, С. Нильсен, Е. Путлей, А. Юнг:
Изготовление и свойства HgTe и смешанных кристаллов HgTe–CdTe . . 325

С. Видергорн, Г. Дрикамер:
Влияние давления на ультрафиолетовые спектры некоторых полициклических ароматических кристаллов 330

Письма в Редакцию:
К. Маеда: Одномерные конфигурационные координатные кривые для центров люминесценции 335

М. Арвин, Р. Типсорд: Коэффициент Холла дисилицида молибдена . . . 336

Рецензия на книгу
« Физика упругости резины », 2-ое издание 338

Band 9 Nu. 1

A. A. Maradudin:
Schraubenversetzungen und diskrete elastische Theorie . . . . . .

G. A. Alers, J. R. Neighbours und H. Sato:
Die Abhängigkeit der Schallgeschwindigkeit und der Dämpfung von der Magnetisierungsrichtung in Nickel bei hohen Feldstärken 21

G. L. Pearson und W. L. Feldman:
Eine Pulvermethode zur Sichtbarmachung der ferroelektrischen Domänenstruktur . . 28

D. G. Thomas:
Die Diffusion und Ausscheidung von Indium in Zinkoxyd 31

M. J. Stephen und A. B. Lidard:
Der Faradayeffekt in Halbleitern 43

VOL. 13 1960

page

H. E. FARNSWORTH und J. TUUL:
Der Einfluss der Dichte der Gitterfehlstellen auf die Chemisorption und die Oxydbildung auf einer reinen (100)-Fläche des Nickels nach Beugungsversuchen mit Elektronen niedriger Energie 48

H. E. FARNSWORTH, R. E. SCHLIER und J. TUUL:
Versuche mit der Brechung von Elektronen niedriger Energie an adsorbiertem Wasserstoff und an Nickelhydroxyd das bei 25°C auf der (100)-Ebene eines Nickelkristalls gebildet wurde 57

F. A. TRUMBORE, C. R. ISENBERG und E. M. PORBANSKY:
Über die Temperaturabhängigkeit des Verteilungskoeffizienten. Die feste Löslichkeit von Zinn in Silizium und Germanium 60

WERNER KÄNZIG und T. O. WOODRUFF:
Die Elektronenstruktur eines H-Zentrums 70

Briefe an die Herausgeber:
E. CORENZWIT: Supraleitfähigkeit von Nb_3Al 93

M. C. STEELE: 93

Band 9 Nu. 2

T. P. HOAR und N. F. MOTT:
Ein Modell für die Bildung von porösen anodischen Oxydschichten auf Aluminium . 97

R. E. SCHMUNK und C. S. SMITH:
Die Ableitungen der elastischen Konstanten von Aluminium und Magnesium nach dem Druck 100

H. C. NASH und C. S. SMITH:
Die elastischen Konstanten von Lithiumeinkristallen. 113

R. T. BATE, R. K. WILLARDSON und A. C. BEER:
Transversale magnetische Widerstandsänderung und Halleffekt in InSb vom *n*-Typ. . . 119

H. EHRENREICH:
Der Transport von Elektronen in InSb 129

N. T. MELAMED:
Energietransport in ZnS:Cu:In-Phosphoren 149

T. RISTE, K. BLINOWSKI und J. JANIK:
Die Streuung von Neutronen in Magnetit infolge Schwankungen der Spinorientierung . 153

G. H. JONKER:
Eine Analyse der Halbleitereigenschaften von Kobaltferriten 165

A. WOLD und R. J. ARNOTT:
Herstellung und kristallographische Eigenschaften der Systeme $LaMn_{1-x}Mn_xO_{3+\lambda}$ und $LaMn_{1-x}Ni_xO_{3+\lambda}$ 176

I. BROSER und R. BROSER-WARMINSKY:
Answer on remarks of M. BALKANSKI about the paper: "On the mechanism of energy transport in Cadmiumsulfid crystals" 181

Briefe an die Herausgeber:
E. BRIAN SMITH: Phasenkoexistenz und Hysterese 182

A. GIANSOLDATI, J. O. LINDE und G. BORELIUS: Die antiferromagnetische Umwandlung der Legierung AuMn 183

Band 9 Nu. 3/4

page

H. A. KLASENS:
Die Temperaturabhängigkeit der Fluoreszenz von Photoleitern 185

S. W. BARBER und B. MARTIN:
Eine Funktion für die Wärmekapazität isotroper Festkörper bei Dispersion der Schallgeschwindigkeit 198

H. J. HROSTOWSKI und R. H. KAISER:
Die Löslichkeit des Sauerstoffs in Silizium 214

F. MOSER, N. R. NAIL und F. URBACH:
Untersuchungen über die optische Absorption der Volumenschwärzung grosser Silberchloridkristalle 217

S. GELLER und M. A. GILLEO:
Die Auswirkung der Dispersionskorrekturen auf die verfeinerte Bestimmung der Yttrium-Eisen-Granatstruktur 235

R. S. KNOX:
Exzitonenzustände in den Kristallen der Edelgase 238

R. S. KNOX:
Exzitonenzustände in kristallinem Argon 265

W. F. LOVE:
Halleffekt in Kupfer und Cu_3Au bei tiefen Temperaturen 281

H. MEYER und P. L. SMITH:
Die Wärmekapazität von sieben Äthylsulfaten der seltenen Erden bei tiefen Temperaturen 285

H. MEYER:
Thermische und magnetische Eigenschaften von Praseodymium-Äthylsulphat unterhalb 1°K 296

R. J. WEISS und A. S. MAE[illegible]TTA:
Die Spinabhängigkeit des elektrischen Widerstandes magnetischer Metalle . . . 302

G. J. WOLGA und M. W. P. STRANDBERG:
Hyperfeinwechselwirkungen in den paramagnetischen Resonanzspektren von LiF und NaF 309

S. ZWERDLING, B. LAX, K. J. BUTTON und LAURA M. ROTH:
Magneto-Absorption von Oszillatoren in Galliumantimonid JA-1149 320

W. D. LAWSON, S. NIELSEN, E. H. PUTLEY und A. S. YOUNG:
Herstellung und Eigenschaften von AgTe und der Mischkristalle HgTe–CdTe . . . 325

S. WIEDERHORN und H. G. DRICKAMER:
Die Druckabhängigkeit der Spektren im nahen ultravioletten Gebeit bei Kristallen aromatischer Ringverbindungen 330

Briefe an die Herausgeber:
K. MAEDA: Eindimensionale Konfigurationskoordinatenkurven für Farbzentren . . 335
M. J. ARVIN und R. F. TIPSORD: Der Hallkoeffizient von Molybdändisilizid . . . 336

Büchbesprechung:
L. R. G. TRELOAR: „The Physics of Rubber Elasticity", Second Edition, Clarendon Press, Oxford, 1958, 242 Seiten, 40s. (Besprochen von B. H. ZIMM und R. W. KILB) . . 338

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 13. pp. 177–186. Printed in Great Britain.

p–i–n JUNCTION IN THE ANODIC OXIDE FILM OF TANTALUM

Y. SASAKI

Research Laboratory, Nippon Electric Company Ltd., Tokyo, Japan

(*Received* 25 *February* 1959; *revised* 28 *July* 1959)

VOL. 13 1960

Abstract—For the purpose of clarifying the structure of anodic oxide film of tantalum, photo-effects and variation of capacity with applied bias potential were studied for relatively thin anodic oxide films. From the results it is concluded that a *p–i–n* junction is formed in the oxide film. The construction of junction is as follows. In contact with the base metal there are very thin *n*-layers, and on the layers there lie nearly stoichiometric *i*-layers. The surface *p*-layers lying above *i*-layers are also very thin. The thickness of the *i*-layers is proportional to the anodization voltage, and the thickness of the extrinsic layers is nearly independent of the anodization voltage and may be in the range of 20–50 Å. The donor levels consist of excess tantalum atoms and the acceptor levels consist of excess oxygen atoms (or defects of tantalum atoms) as well as adsorbed oxygen on the oxide surface. Due to nearly amorphous structure of the oxide, the recombination rate of injected carriers is high in the junction. Coincidence between the theory of *p–i–n* junction and the experimental results is quite good.

INTRODUCTION

THE capacitor action[1] and the rectifying action[2–5] of anodic oxide film of tantalum are well known. Lately we found that the solid Ta capacitor,[6] obtained by evaporation of some semiconductors such as germanium upon anodized tantalum foil, has nearly the same capacity as that of the electrolytic Ta capacitor generally used. Due to the fact that the capacitor contains no electrolyte, it can be used in the wide temperature range, viz. −200°C to +200°C. The other properties of the capacitor, such as the rectifying action, the break down voltage and the variation of capacity with applied bias potential, are also nearly the same as those of the electrolytic Ta capacitor. This suggests that these properties come from the anodic oxide film, but not from the contact zone between the oxide surface and the electrolyte. In the literature we find tentative explanations of the mechanism of rectification of the anodic oxide film of Ta, viz. the existence of *p–n* junction,[2] the Mott–Schottky theory,[3] the existence of trapped ionic space charge[4] and local rectification at isolated area due to the imperfections of the oxide film.[5]

With the purpose of presenting a view that the rectifying action takes place as the natural consequence of the inner structure of the film, studies were carried out on the photo-effects and on the variation of capacity with applied bias potential. These phenomena cannot be explained with the simple model[1] that the electrolytic Ta capacitor consists of Ta base metal, dielectrics of Ta_2O_5 and the electrolyte. The obtained results suggest strongly the existence of the space charge in the oxide due to excess Ta atoms (donor levels) at the innermost region and also to the excess oxygen atoms or defects of Ta atoms (acceptor levels) at the surface region. Between the extrinsic regions there is a nearly stoichiometric region, whose thickness is proportional to the anodization voltage.

In our experiments, relatively thin films of anodic oxide were used (usually less than 500 Å). The photo-voltaic effects and the variation of capacity were observed more clearly for thinner films than for thicker films. This effect of film thickness is reasonable, because the extrinsic regions are very thin, the intrinsic region is relatively thick and the recombination rate of injected carriers in these regions may be high due to nearly amorphous structure of the oxide film.

EXPERIMENTS

(a) *Preparation of samples*

Tantalum foil of 0·5 mil thickness (product of the Fansteel Metallurgical Co., purity>99·9 per cent), is rinsed in acetone and anodized at 10 V. After anodization, the foil is dipped into 20 per cent hydrofluoric acid in order to remove the anodic oxide film. The foil is again washed and anodized in ammonium borate solution for 24 hr. The resulting oxide film has small leakage current at the bias of anodization voltage ($<0{\cdot}01\mu A/cm^2$); its surface, examined with an electron microscope, is as smooth as that of chemically polished Ta foil.[7] The anomaly observed by VERMILYEA[8] and YOUNG[9] for Ta foils having rough surface was not observed. For the purpose of comparison, chemically polished foils were also used in some of the following experiments. The results obtained were almost the same in both cases.

(b) *Photo-voltage*

Ta foils, anodized at 4–40 V, are immersed in electrolytes contained in a quartz vessel with a platinum reference electrode, and irradiated with ultra-violet light of Hg vapor lamp. The pH values of the electrolytes are varied in the range 1–13 by using H_2SO_4, NaOH and ammonium borate. The absorption coefficients of the solutions for the ultra-violet light are very small and can be neglected. The foil shows photo-current (I_L)[10–14] and photo-voltage (V_{photo})[11–12]. A typical example of the relation between I_L and the applied bias potential ($\delta\phi$) is shown in Fig. 1 for a Ta foil anodized at 10 V. According to the notations generally used in the *p–n* junction theory, $\delta\phi$ is taken to be negative when the potential is applied

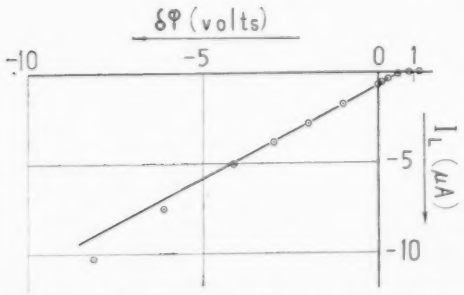


FIG. 1. Photo-effect of anodized Ta foil caused by irradiation of ultra-violet light.

in the direction of hard flow (the Ta electrode positive and the Pt electrode negative). The positive $\delta\phi$ in the figure represents photo-voltage. The observed value of I_L at a given $\delta\phi$ is proportional to the intensity of irradiation. Thus I_L, under the irradiation of ultra-violet light having intensity L (arbitrary unit), can approximately be expressed by

$$I_L = L(\phi_0+\delta\phi)/K, \qquad (1)$$

where the constants, K and ϕ_0, are independent of pH of the solution, and the value of K (resistance of the oxide at unit intensity of irradiation) increases with the increase of thickness of the oxide film.

By irradiating the foil with monochromatic ultra-violet light from Beckman Model *D–U* Photometer, the threshold wave length of the photo-current is found to be (3400±100)Å, corresponding to ca. 3·6 eV. The threshold wave length of the photo-voltage is nearly the same (about 3000 Å, i.e. 4·1 eV). From the directions of the photo-voltage and photo-current, photo-emission of electrons from the Ta metal should be excluded. Further, the facts (i) that the slope of plots of I_L versus $h\nu$ of the impinging photons is very steep (see Fig. 11) and (ii) that the absorption coefficients of the oxide for ultra-violet light having wave length less than 2800 Å are more than 10^5 cm^{-1}, suggest that the energy of 3·6–4·1 eV is attributed to the energy gap of the oxide.[15]

The measurements of the voltage, which was observed between the Ta and the Pt electrodes with and without irradiation of ultra-violet light, were made with a vacuum tube voltmeter, whose input impedance is greater than $10^{13}\Omega$. When the Ta foil—from which the oxide film formed by the action of air is removed by hydrofluoric acid just before the measurement—is immersed in an electrolyte (1N H_2SO_4), the Pt electrode shows positive potential (V_{react}) of about 0·8 V relative to the Ta electrode due to electrochemical reaction between the Ta metal and the electrolyte.[16] V_{photo} is not observed on irradiation. When the circuit between the two electrodes is closed, electric current flows between the electrodes and reaction proceeds forming oxide layers on the Ta metal. Then V_{react} becomes smaller, and small V_{photo} appears. In a few days V_{react} reduces nearly to zero and V_{photo} of about 1 V is observed.

VOL. 13 1960

With Ta foils, anodized at a voltage larger than 1 V, V_{react} is always zero. The open circuit V_{photo} of the anodized Ta foils immersed in ammonium borate solution (pH = 6·0), was measured. V_{photo} of a foil, anodized at 10 V, at time t (V_t) are given in Fig. 2 as a function of the duration of irradiation

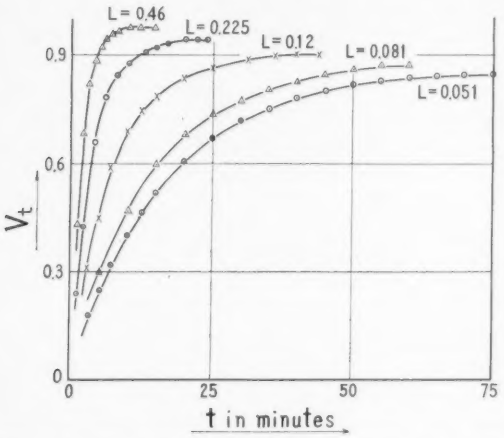


FIG. 2. Variation of photo-voltage (in V) with time at various intensity of light (arbitrary unit) at 24°C.

for various L. The magnitude of L, expressed in terms of the photo-current of the Ta foil at $\delta\phi = -2$ V, was 0·005 μA $\sim$ 3 μA/cm^2. V_t at constant L increases with time and finally reaches a maximum

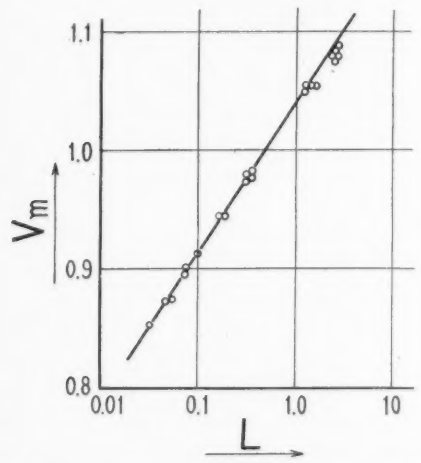


FIG. 3. Relation between V_m (in V) and L (in log scale) at 20°C.

saturation value V_m. The relation between V_m and $\log_{10}L$ is given in Fig. 3. The observed maximum values of V_m were 1·2–1·3 V for Ta foils anodized at 5–20 V. The slope of the line in Fig. 3 is about $2{\cdot}3\times 2kT/q$ (= 0·12 V at room temperature), where k is Boltzmann's constant, T is absolute temperature and q is the electronic charge. The temperature dependence of V_m at constant L is given in Fig. 4.

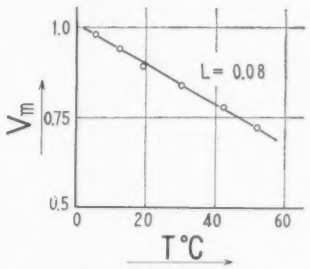


FIG. 4. Temperature dependence of V_m at L = 0·08.

When the irradiation is carried out without external application of negative $\delta\phi$, V_m at a given L and T does not change with the duration of irradiation of large intensity for 100 hr. After the irradiation, the thickness of the oxide film is determined by means of interference colour and capacity of the foil, and is found to remain almost unchanged compared with the original value. It is known that the abnormal growth of the oxide film thickness takes place when anodized Ta foil is irradiated with ultra-violet light at the bias of large negative $\delta\phi$[12–14], especially when $\delta\phi$ is nearly equal to the anodization voltage. In the experiments of photo-current at $\delta\phi \geqq 0$ and of open circuit photo-voltage, however, the irradiation has always been carried out without externally applied field, and the film growth, i.e. the reaction between the Ta metal and the electrolyte through the oxide film, does not take place or, at least, can be neglected.

The measurement of photo-voltage was carried out in the pH range 1–13. It was found that for all solutions the relations between V_t and t show little difference and the slope of V_m versus $\ln L$ is always $2kT/q$. The values of V_m at constant L and T remain unchanged in the pH range 1–10 (Fig. 5). When pH exceeds 10, V_m decreases with the increase of pH.

It is clear that the anodic oxide film is responsible to the photo-voltaic effects. The facts that, (i) when $\delta\phi$ is positive, no noticeable reaction takes place between the metal and the electrolyte, (ii) the rectifying action of anodized Ta foil immersed in an electrolyte is nearly the same as that of the Ta foil, on which thin germanium film is evaporated as an electrode and (iii) there is little correlation between the photo-effects and pH of the solution, suggest that I_L at positive $\delta\phi$ and V_{photo} are due to the inner structure of the oxide film, viz. space charge existing in the oxide film,† and that the electrolyte does not play an important role.

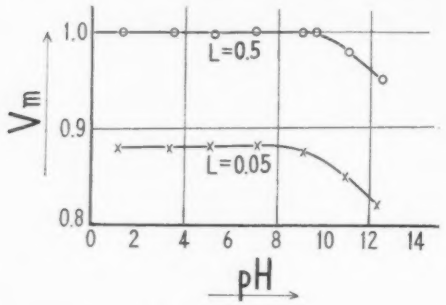


FIG. 5. Variation of V_m with pH of the electrolyte at 20°C.

In the anodic oxide film of Ta, the mobility of electron will be low and the recombination rate of injected carriers will be high, because the oxide has ionic and nearly amorphous[21] structure. The observed slope of $2kT/q$ (Fig. 3) suggests that the oxide film has a p–i–n, n^*–n (where n^* is heavily doped n-type Ta_2O_5)‡ or similar structure, where relatively fast recombination of injected carriers takes place in the transition region of the junction. It is known that the observed maximum V_{photo} of p–n junction of Ge or Si at room temperature is about one half of E_g of the semiconductor used.[22] The maximum values of V_{photo} of Ta foils anodized at 0–20 V (that is, the thickness of the oxide film is in the range 100–500 Å) are 1–1·3 V. Thus the difference between the electrostatic potential of the innermost oxide layers and that of the surface oxide layers in thermal equilibrium may be greater than 3 eV, and it seems more reasonable to assume p–i–n junction than to assume n^*–n junction. If such a junction really exists, variation of capacity with applied bias potential and rectifying action are expected for the anodic oxide film.

(c) *Variation of capacity with applied potential*

The variation of capacity C of electrolytic Ta capacitors, anodized at 6–40 V, as well as that of the solid Ta capacitors with applied bias potential were measured with a.c. bridge. The fluctuation of the observed value of capacitance at $\delta\phi$, $C_{\delta\phi}$, lies within 0·02 per cent. An example of the relations between $\delta\phi$ and $C_{\delta\phi}$ and between $\delta\phi$ and electric current $I_{\delta\phi}$ is shown in Fig. 6 for an electrolytic Ta capacitor anodized at 10 V. When $\delta\phi$ is negative, the decrease in capacity with $\delta\phi$ is relatively small. If $\delta\phi$ is positive, especially when $\delta\phi \geqq 0{\cdot}9$ V, large increases in $C_{\delta\phi}$ and $I_{\delta\phi}$ are observed.

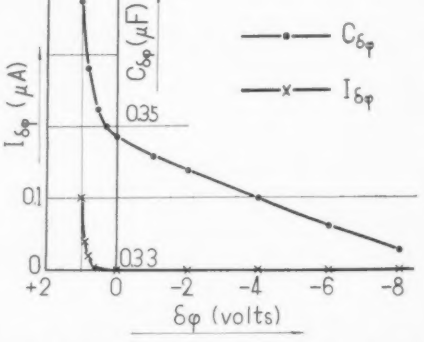


FIG. 6. Variation of capacity $C_{\delta\phi}$/cm² and of electric current $I_{\delta\phi}$/cm² of an electrolytic Ta capacitor anodized at 10 V with applied bias potential ($\delta\phi$) at 20°C.

Ta foil generally used is known[23] to contain non-metallic inclusions of appreciable size. VERMILYEA[5] has reported that when Ta foils having thick oxide films (1750 Å thick) are used, rectification occurs principally at singularities in the surface, presumably at such inclusions. The same experiment was carried out for thinner oxide films

† It was shown by MOTT[17] that the photo-voltage is observed only when there is space charge in the semiconductor and when the semiconductor is irradiated by the light of which energy is greater than the band gap. Existence of space charge in anodic oxide film of Ta has been suggested by HARING[4], APKER and TAFT[11], DEWALD[18–19] and by YOUNG[20].

‡ The n^*–n structure has been suggested by several authors.[4,11,18,19]

by using mercury wetted copper contact of small area. It is found that the rectifying action is clear for the thin oxide films, especially when the anodization voltage is less than 10 V. Forward current in the range of $1\text{–}3\times10^{-9}$ A at $\delta\phi=+2{\cdot}6$ V and backward current of less than 10^{-10} A at $\delta\phi=-2{\cdot}6$ V are observed for Ta foils anodized at 6 V at any point on the oxide surface. The resistance in the easy flow direction increases with the increase of film thickness, and the rectification becomes indistinct for Ta foils anodized at a voltage greater than 20 V. The same result was obtained for chemically polished Ta foils. If we assume the existence of p–i–n junction in the oxide film, the large increase in $I_{\delta\phi}$ and in $C_{\delta\phi}$ at positive $\delta\phi$ can easily be explained by the conductivity modulation and by the addition of diffusion capacitance caused by carrier injection. The observed large resistance in the easy flow direction, which increases with the increase in film thickness, is also explainable, because low mobility of current carriers and high recombination rate of injected carriers are expected in the i-region, whose thickness increases with the increase of anodization voltage.

The variations in $C_{\delta\phi}$ and $I_{\delta\phi}$ of the solid Ta capacitor are also found to be nearly the same as those of electrolytic Ta capacitor. When $\delta\phi\leqq0$, the change of capacity of the two capacitors is the same. When $\delta\phi>0$ and its absolute value is small and the duration of observation is short, the capacity of the solid Ta capacitor increases in the same way as shown in Fig. 6. The capacity, however, decreases unsteadily with the lapse of time due to destruction of very small patches of the oxide film. When $\delta\phi>5$ V, the solid Ta capacitor undergoes often complete shortening. So the value of capacity of the solid Ta capacitor for $\delta\phi>0$ cannot be determined exactly.

DISCUSSION

(a) *p–i–n junction model*

There are different views on the structure of anodic oxide film of Ta, especially on the existence of space charge[4, 5, 11, 18–20, 24–26] in the oxide due to concentration gradient of excess tantalum ions or of excess oxygen ions. From the experimental results of the photo-effects and of the variation of capacity with $\delta\phi$, p–i–n junction is expected to be formed in the anodic oxide film.

CABRERA and MOTT[27] have suggested that local thermodynamic equilibrium exists at the metal–oxide interface and at the oxide–oxygen interface. If the oxide can dissolve excess metal atoms, the oxide layers adjacent to the metal would become n-type semiconductor, and if the oxide can dissolve oxygen, the surface oxide layers would become defect type, i.e. p-type. As Ta_2O_5 is known to be the semiconductor of n-type having very large specific resistivity,[28] it may be expected that the n-region exists adjacent to the metal. During the course of anodization, however, the concentration of adsorbed oxygen ions at the oxide surface is very high, and it might be assumed that the surface region is p-type. The structure of the oxide film would be determined by the behavior of carriers of anodizing current in the course of anodization. Three theories have been proposed to explain the experimental results[7, 12, 24–26, 29–30] of anodization, viz. (i) theory by CABRERA and MOTT[27], (ii) by DEWALD[19] and (iii) by BEAN *et al.*[31]. Assuming that the current carriers in the anodization process are interstitial Ta ions, the last two theories predict the presence of n-layers adjacent to the base metal due to excess Ta atoms and that next to the n-layers there lie nearly stoichiometric oxide layers. The surface p-layers might be formed if the anodizing current is also carried by oxygen ions or if a large number of oxygen atoms or ions are adsorbed on the oxide surface.

In the intermediate region of the oxide film between base metal and oxide surface, the concentration of the impurity levels may be scarcely influenced by the presence of base metal and of electrolyte.[27] The concentration would be very low because Ta_2O_5 is generally a good insulator. Thus, we will assume that a p–i–n junction exists in the anodic oxide film, and will discuss the experimental results in the light of p–i–n junction.

(b) *Capacity of p–i–n junction**

For brevity of calculation of junction capacity, the distribution of effective impurity concentration (N_d-N_a) in the oxide film is assumed as follows (Fig. 7):

(i) $d(N_d-N_a)/dx=a$ in the n- and p-regions $(D/2>x>d$ and $-d>x>-D/2)$,

* The calculation is analogous to that of p–n junction capacity reported by SHOCKLEY[32].

(ii) $N_d - N_a = 0$ in the i-region ($d \geqq x \geqq -d$), where a is a constant, D = the thickness of oxide film, $2d$ = the thickness of i-region, and N_d and N_a are the concentrations of donors and of acceptors, respectively. For one dimensional case, Poisson's equation is

$$d^2\Psi/dx^2 = -4\pi\rho/\kappa, \tag{2}$$

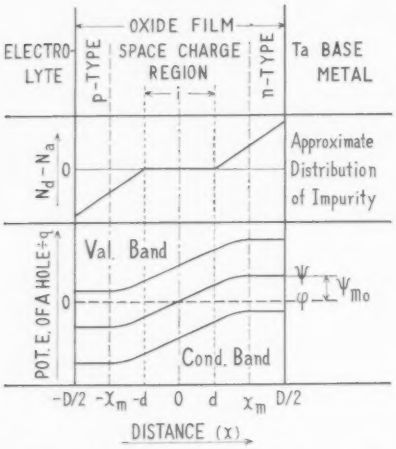


FIG. 7. Schematic diagram of $N_d - N_a$ and Ψ in the anodic oxide film of Ta.

where ρ is the charge density and κ is the dielectric constant of Ta_2O_5. The electrostatic potential Ψ is determined with respect to the Fermi potential ϕ in thermal equilibrium. We assume that all the impurity levels are singly ionized in the space charge region, and to the first approximation, when $\delta\phi \leqq 0$, we neglect the electron and hole space charge in the transition region. Using the boundary conditions that $d\Psi/dx = 0$ at $x = \pm x_m$ (viz. neutral extrinsic oxides begin at $x = \pm x_m$), we obtain the following equations.

$$\Psi_1 = 2\pi qa\{3(x_m - d)^2 x - (x-d)^3\}/3\kappa, \tag{3}$$

$$\Psi_2 = 2\pi qa(x_m - d)^2 x/\kappa, \tag{4}$$

$$\Psi_3 = 2\pi qa\{3(x_m - d)^2 x - (x+d)^3\}/3\kappa, \tag{5}$$

where the zero of potential is the value at $x = 0$, a condition required by the symmetry between $+x$ and $-x$ of the assumptions (i) and (ii). Ψ_1, Ψ_2 and Ψ_3 are the electrostatic potentials in the n-, i- and p- regions, respectively. Ψ is also shown schematically in Fig. 7. The capacity of the junction per unit area, C, is given by

$$C = \kappa/4\pi(2x_m). \tag{6}$$

Let Ψ_{m_0} be the value of Ψ_1 at $x = x_m$ for which $\delta\phi = 0$, then the value of Ψ_1 at $x = x_m$ when $\delta\phi$ is applied is equal to $(\Psi_{m_0} - \delta\phi/2)$. Thus we get

$$2\Psi_{m_0} - \delta\phi = 4\pi qa\{2(x_m)^3_{\delta\phi} - 3(x_m)^2_{\delta\phi} d + d^3\}/3\kappa, \tag{7}$$

where $(x_m)_{\delta\phi}$ is the value of x_m when $\delta\phi$ is applied.

(c) *Thickness of the intrinsic and extrinsic regions*

It is well known that the thickness of the anodic oxide film of Ta increases with the anodization voltage at a constant rate of 10–20Å/V. The values of x_m when anodization voltage V_0 is applied, $(x_m)_{V_0}$, are calculated from the values of C_{V_0} by means of equation (6) under the assumptions that $\kappa = 23$[12, 33] and that the roughness factor of the oxide film is unity. The calculated values of

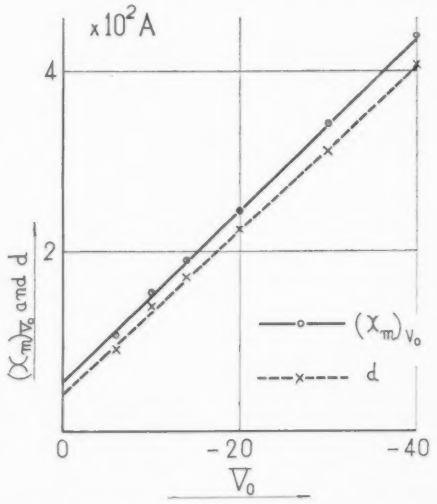


FIG. 8. Variation of $(x_m)_{V_0}$ and of d with anodization voltage (V_0) in V. The roughness factor of the oxide film is assumed to be unity.

$(x_m)_{V_0}$ are plotted in Fig. 8 against V_0. It is seen that $(x_m)_{V_0}$ is approximately expressed by

$$2(x_m)_{V_0} = (19V_0 + 110)\text{Å}. \tag{8a}$$

The thickness of the i-region can also be calculated by means of equation (7) from the experimental results of the variation of capacity of the

oxide film at relatively low applied potential. From the large value of capacity of anodized Ta, the impurity concentration at x_m should be very large. Thus we assume Ψ_{m_0} is nearly equal to $E_g/2$, i.e. 2 V. The calculated values of d of the oxide films are also plotted in Fig. 8. The relation between d and V_0 is

$$2d = (18{\cdot}5V_0 + 75)\text{Å}. \qquad (8b)$$

From equations (8a) and (8b) it is concluded that the thickness of the i-region increases in proportion to the anodization voltage and the values of $\{(x_m)_{V_0} - d\}$ are almost independent of the anodization voltage.

From equation (7) we see that the relation between

$$\{2(x_m)_{\delta\phi}{}^3 - 3(x_m)_{\delta\phi}{}^2 d + d^3\}$$

and $\delta\phi$ should be represented by a straight line with the slope $-3\kappa/4\pi qa$, and the extrapolation of

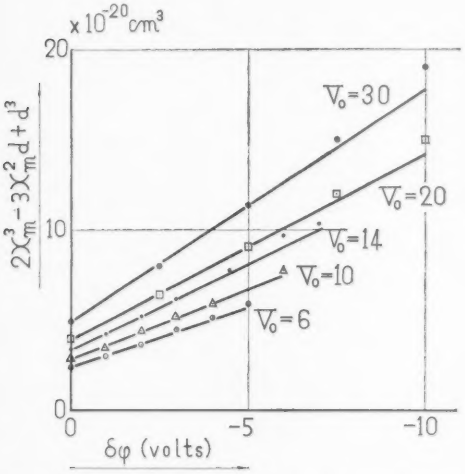


FIG. 9. Variation of $\{2(x_m)_{\delta\phi}{}^3 - 3(x_m)_{\delta\phi}{}^2 d + d^3\}$ of the electrolytic Ta capacitors, anodized at 6–30 V, with $\delta\phi$.

the line should intersect the abscissa at $2\Psi_{m_0}$. In Fig. 9 the values of

$$\{2(x_m)_{\delta\phi}{}^3 - 3(x_m)_{\delta\phi}{}^2 d + d^3\}$$

are plotted against $\delta\phi$ for the foils anodized at various V_0. When $\delta\phi \leqq 0$, i.e. when no carriers are injected into i-region, we find that the experimental results are in good agreement with equation (7). In Table 1 we summarize the values of $(x_m)_{V_0}$,

Table 1. *The anodization voltage* (V_0), *the effective impurity concentration gradient* (a), $(x_m)_{V_0} - d$, *and the impurity concentration at* $x = (x_m)_{V_0}$, N_m, *are shown for electrolytic Ta capacitor anodized at* 6–40 V.

V_0 (V)	a $\times 10^{-27}$/cm⁴	$(x_m)_{V_0} - d$ (Å)	$N_m \times 10^{-20}$ at $x = (x_m)_{V_0}$
6	5·4	14·8	8·0
10	4·8	16·9	8·1
14	4·0	19·6	7·8
20	3·6	21·4	7·7
30	2·5	27	6·8
40	2·0	30·5	6·1

impurity concentration gradient (a), which is calculated from the slope of the lines in Fig. 9, and impurity concentration at $x = (x_m)_{V_0}$, N_m, for the oxide films anodized at 6–40 V.

It is interesting to examine the values listed in the Table 1 in the light of anodization mechanism. Under the influence of externally applied electric field, E, the number of ions removing from interstitial positions to the neighbouring interstitial positions, n/cm²/sec, is given by[19]

$$n = \text{const. } N \exp[\{\lambda qE - U\}/kT], \qquad (9)$$

where N is the impurity concentration and λ is the half distance between interstitial positions. U is the activation energy for motions of interstitial ions between interstitial positions in the absence of electric field, and is estimated to be about 1·5 eV [19, 29, 31]. Therefore, the diffusion of interstitial ions in the extrinsic regions is negligible, though the gradient is very steep. By applying equation (9) for the i-region of the p–i–n junction, it is readily seen that the value of electric field in the i-region at the bias of anodization voltage, E_{V_0}, should be independent of the anodization voltage, because the values of N, λ and U in the i-region can reasonably be assumed to be independent of the thickness of the oxide film. The value of E_{V_0} is

$$E_{V_0} \equiv -d(\Psi_2)_{V_0}/dx = -2\pi qa\{(x_m)_{V_0} - d\}^2/\kappa, \qquad (10)$$

where $(\Psi_2)_{V_0}$ is the value of Ψ_2 at $\delta\phi = V_0$. The values of E_{V_0}, obtained by using the values of $\{(x_m)_{V_0} - d\}$ and a are 4–7 $\times 10^6$ V/cm for the

oxide film of various thickness. The value agrees fairly well with equation (8a).

(d) *The surface p-layers*

Some experiments were carried out to study the surface of the oxide film. Ta foils, anodized at 20 V, are heated for 2 hr in high vacuum or in oxygen atmosphere of 760 mm Hg. The temperature range of the heat-treatment is 100°C–500°C. Immersed in ammonium borate solution, the capacity of the heat-treated foil (C_h) is measured at $\delta\phi = 0$. The results obtained are shown in Fig. 10. When the foils are heated in vacuum, C_h decreases with the increase of temperature above

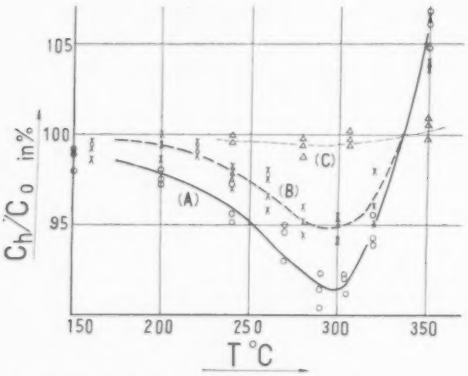


FIG. 10. Variation of capacitance with heat-treatment. Curve *A* (Ta foils etched by the usual method) and curve *B* (chemically polished Ta foils) are for the heat-treatment in vacuum, while curve *C* is for heat-treatment in oxygen atmosphere.

200°C, reaches a minimum at about 300°C, and increases appreciably above 350°C. The capacitance at the minimum point is 5–8 per cent below that of the not-heat-treated foil (C_0). When heated in oxygen, C_h remains almost unchanged in the range 100°C–500°C.

On applying −20 V to the heated foil, charging current of the capacitance plus anodizing current flows for a short time, then the capacitance becomes nearly equal to the original value in 1–2 min. Anodizing current is small for the foils heated in vacuum when the temperature is below 200°C. The current increases appreciably above 220°C, and greatly above 350°C. No bubble is observed on the oxide surface in the anodization process. The increase of anodizing current is small for the foils heated in oxygen, especially when the temperature is below 300°C.

As was reported by VERMILYEA[34], the large increase of capacity, which takes place above 350°C can be explained by the diffusion of oxygen atoms from the oxide into the base metal or of Ta atoms from the base metal into the oxide. The small decrease of capacity and small increase of current in the temperature range 200°–300°C, however, cannot be explained by the diffusion, since the diffusion would give rise to the increase in capacity. The variation of C_h at 250°–280°C, caused by heating cycles of the foil in vacuum and in oxygen atmosphere alternatively, is reproducible. Effect of oxygen pressure during the heat-treatment at 280°C was measured in the pressure range 10^{-5} −800 mm Hg. C_h increases with the increase of oxygen pressure.

ISHIKAWA[35] reported that in the case of BaO, *p–n* transition was caused by sorption and liberation of oxygen atoms and the sorption took place at relatively low temperature (about 300°C). If we assume the existence of very thin surface *p*-layers due to adsorbed oxygen atoms or ions in and on the oxide surface, these atoms would be liberated or adsorbed by heating the foil in vacuum or in oxygen atmosphere at relatively low temperature. The reduction of excess oxygen would make the *p*-layers less *p*-type or nearly intrinsic and would cause the decrease of capacity. The reduction of capacity of 5–8 per cent corresponds to the increase of the thickness of *i*-layers by 20–30 Å. This value coincides with the thickness of surface *p*-layers calculated from the variation of capacity of anodized Ta foil with $\delta\phi$.

It was observed that electric current of electrolytic Ta capacitor in the direction of hard flow increases greatly for a short time by the preceding applications of positive $\delta\phi$, and that solid electrolytic Ta capacitor is easily destroyed by the application of positive $\delta\phi$. The phenomena are also explainable by the presence of a large number of negative ions on the oxide surface, because the adsorbed ions may be liberated from the oxide surface by the application of positive $\delta\phi$.

(e) *Photo-effects*

Photo-voltaic effect with semiconductors is observed when the following conditions are satisfied: (i) space charge exists in the semiconductors

in thermal equilibrium and (ii) bulk life time of the minority carriers generated by the irradiation is not too small compared with the transit time for crossing the transition region. The mobility of electron μ in ionic crystals, such as BaO and TiO_2, is known to be in the order of 1 cm^2/V sec.[36] The life time of the excess minority carriers in the neutral p-region, τ, is supposed to be very small, because the oxide is nearly amorphous and the concentration of acceptors in the region is very high. Thus the diffusion length ($=\sqrt{(D\tau)}$, where $D = \mu kT/q$) may be smaller than the thickness of the neutral p-region, and few excess carriers generated in the p-region would get to the transition region. The carriers generated in the transition region, however, move fast under the influence of the large electric field. Thus we assume that only the electrons and the holes generated in the transition region can reach the n- and the p-region of the oxide, respectively, and give rise to the photo-effects.

As is seen in Fig. 1, the photo-current of anodized Ta foil increases with the increase of $-\delta\phi$. The result may be explained qualitatively in the following processes: (a) the widening of the space charge region by the applied potential, that is, the reduction of thickness of the neutral p-region and (b) recombination of injected carriers in the space charge region. Thus the photo-current is given by

$$I_L = \alpha L, \tag{11}$$

where α is an apparent quantum efficiency of an impinging photon and is a function of $\delta\phi$. Variation of α at $\delta\phi = 0$ with the wave length of incident light is measured for a Ta foil anodized at 10 V (Fig. 11). The value of 1/50 may be reasonable for such a high recombinating junction as anodic oxide film.

The open circuit photo-voltage of anodic oxide film can be deduced from the p–i–n junction model. In the steady state, the generation current due to photon absorption and the forward current I_f should cancel each other.[22] The electric field in the oxide under the steady state is

$$(2\Psi_{m0} - V_m)/2(x_m)V_m.$$

The change of V_m with L is not large (Fig. 3), and

$$(2\Psi_{m0} - V_m)$$

is approximated to be constant. Thus we get

$$I_f \propto L. \tag{12}$$

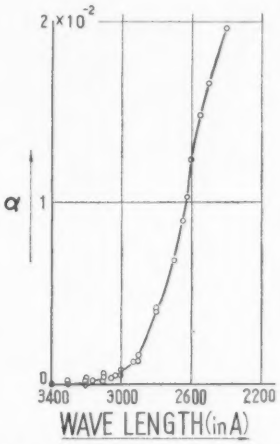


FIG. 11. Variation of apparent quantum efficiency α with wave length of the incident light.

I_f (at the bias of V_m) of p–i–n junctions of Ge,[37, 38] Si[22] and SiC[39], with which the recombination of the injected carriers takes place mainly in the i-region, is expressed by the following equation,

$$I_f = I_s \exp qV_m/AkT, \tag{13}$$

where I_s is the saturation current of the junction and A is approximately two. The forward current of anodized Ta foil, however, does not agree with equation (13). The reason of the disagreement may be as follows. When external forward bias is applied to the oxide film, electric current flows through the external circuit. The current causes (i) ohmic drops in the n- and the p-region and at the oxide–electrolyte interface and (ii) the increase of leakage current by the liberation of adsorbed ions on the oxide surface. In the case of photo-voltage, however, I_f does not flow through the external circuit. Thus equation (13) may hold good for I_f in the case of photo-voltaic effect. Combining equations (12) and (13), we get

$$V_m = \frac{2kT}{q} \ln L + \text{constant}.$$

Thus the relation between V_m and $\ln(L)$ should be represented by a straight line having the slope of

$2kT/q$. This is in good agreement with the photo-experiments shown in Fig. 3.

In the light of the *p–i–n* junction model we can explain quantitatively such phenomena as photo-effects, variation of capacity with $\delta\phi$, which are beyond the reach of simple theory that the electrolytic Ta capacitor consists of the base Ta metal, dielectrics and electrolyte. It follows that the electrolyte is not the necessity to Ta capacitors. The mechanism of solid electrolytic Ta capacitor using no electrolyte can satisfactorily be explained in view of the *p–i–n* junction.

Acknowledgements—The author wishes to thank Dr. M. KOBAYASHI, Director of the Laboratory of Nippon Electric Company Ltd., for his encouragement. He is especially grateful to Dr. Y. ISHIKAWA for his helpful guidance during the course of this work and Mr. E. OKAMOTO for his assistance in the experiments.

Thanks are due to Prof. K. IWASE and Prof. H. TAKAKI of Kyoto University and Dr. J. YAMASHITA for revising this paper.

REFERENCES

1. cf. GÜNTHERSCHULZE A. and BETZ H., *Electrolyt-Kondensatoren* (2. Aufl.). Herbert Cram, Berlin (1952).
2. SCHOLTE J. W. A. and VAN GEEL W. Ch., *Philips Res. Rept.* **8**, 47 (1953).
3. ANDERSON S., *J. Appl. Phys.* **14**, 601 (1943).
4. HARING H. E., *J. Electrochem. Soc.* **99**, 30 (1952).
5. VERMILYEA D. A., *J. Appl. Phys.* **27**, 963 (1956).
6. KOBAYASHI M., *Proc. Electronic Components Conference*, Philadelphia Pa. p. 62, May (1959).
7. VERMILYEA D. A., *Acta Met.* **1**, 282 (1953).
8. VERMILYEA D. A., *Acta Met.* **2**, 476 (1954).
9. YOUNG L., *Acta Met.* **5**, 711 (1957).
10. BÄR W., *Z. Phys.* **115**, 658 (1940).
11. APKER L. and TAFT E. A., *Phys. Rev.* **88**, 58 (1952).
12. YOUNG L., *Trans. Faraday Soc.* **50**, 153 (1954).
13. VERMILYEA D. A., *J. Appl. Phys.* **26**, 489 (1955).
14. VERMILYEA D. A., *J. Electrochem. Soc.* **104**, 212 (1957).
15. MOTT N. F. and GURNEY R. W., *Electronic Processes in Ionic Crystals* p. 104. Oxford University Press (1940).
16. VERMILYEA D. A., *J. Electrochem. Soc.* **101**, 389 (1954).
17. MOTT N. F. and GURNEY R. W., *Electronic Processes in Ionic Crystals* pp. 192–196. Oxford University Press (1940).
18. DEWALD J. F., *Acta Met.* **2**, 340 (1954).
19. DEWALD J. F., *J. Electrochem. Soc.* **102**, 1 (1955).
20. YOUNG L., *Trans. Faraday Soc.* **51**, 1250 (1955).
21. VERMILYEA D. A., *J. Electrochem. Soc.* **102**, 207 (1955).
22. PRINCE M. B., *J. Appl. Phys.* **26**, 534 (1955).
23. VERMILYEA D. A., *J. Electrochem. Soc.* **104**, 542 (1957).
24. VERMILYEA D. A., *Acta Met.* **2**, 340 (1954).
25. VERMILYEA D. A., *Acta Met.* **3**, 106 (1955).
26. VERMILYEA D. A., *J. Electrochem. Soc.* **104**, 140 (1957).
27. CABRERA N. and MOTT N. F., *Reports on Progress in Physics*. Vol. 12, p. 163. Physical Society, London (1949).
28. HARTMANN V. W., *Z. Phys.* **102**, 709 (1936).
29. VERMILYEA D. A., *J. Electrochem. Soc.* **102**, 655 (1955).
30. BRAY A. R., JACOBS P. W. M. and YOUNG L., *Trans. Faraday Soc.* **71**, 405 (1958).
31. BEAN C. P., FISHER J. C. and VERMILYEA D. A., *Phys. Rev.* **101**, 551 (1956).
32. SHOCKLEY W., *Bell Syst. Tech. J.* **28**, 435 (1949).
33. YOUNG L., *Proc. Roy. Soc.* A**244**, 41 (1958).
34. VERMILYEA D. A., *Acta Met.* **5**, 113 (1957).
35. ISHIKAWA Y., SATO T., OKUMURA K. and SASAKI T., *Phys. Rev.* **84**, 371 (1951).
36. *American Institute of Physics Handbook*. pp. V–161, McGraw-Hill, New York (1957).
37. HALL R. N., *Proc. Inst. Radio Engrs.* **40**, 1512 (1952).
38. SAH C. T., NOYCE R. N. and SHOCKLEY W., *Proc. Inst. Radio Engrs.* **45**, 1228 (1957).
39. PATRICK L., *J. Appl. Phys.* **28**, 765 (1957).

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 13. pp. 187–203. Printed in Great Britain.

REACTIONS OF GROUP III ACCEPTORS WITH OXYGEN IN SILICON CRYSTALS

C. S. FULLER, F. H. DOLEIDEN and KATHERINE WOLFSTIRN

Bell Telephone Laboratories, Murray Hill, N.J.

(*Received* 13 *August* 1959)

VOL. 13 1960

Abstract—Hall effect and conductivity results are presented on Si crystals containing from 5×10^{17} to $1{\cdot}6\times10^{18}$ cm^{-3} of oxygen and various dopings of B, Al and Ga before and after reaction at various temperatures (410–1350°C). Evidence in support of the tetrahedral SiO_4 model for the donor formed from oxygen suggested by Kaiser is presented. A large hole–electron effect is noted in the presence of excess acceptor concentration. In addition, specific interactions with the acceptor atoms occur. Although the reactions are complex and are not understood, three kinds are postulated in order to account for the results: (1) The formation of isolated SiO_4 donors such as occurs in undoped Si crystals containing oxygen. (2) The reaction of oxygen at an acceptor atom to form a neutral compound or a "molecular pair". (3) The formation of a donor involving eight oxygen atoms and an acceptor atom. The latter reaction is believed to occur most readily in the case of Al-doped crystals and, in an excess of oxygen, leads to the formation of one donor per Al atom.

1. INTRODUCTION

SOLUTIONS of oxygen in silicon containing $\sim10^{18}$ cm^{-3} oxygen, and no other doping agents, when subjected to temperatures between 400–1000°C undergo changes both in electrical properties and in infrared absorption.[1–5] These changes can be explained[6] qualitatively at least, by assuming a series of consecutive reactions of the type:

$$\begin{aligned} &O+SiO \rightleftarrows SiO_2\\ &O+SiO_2 \rightleftarrows SiO_3\\ &O+SiO \rightleftarrows SiO_4 \rightleftarrows SiO_4^+ + e^- \end{aligned} \qquad (1)$$

where O stands for dissolved oxygen and the silicon–oxygen compounds represent radicals which form by the successive addition of oxygen to a specific Si atom of the lattice. The unit containing four oxygen atoms, first postulated by W. KAISER[4], is believed to be a particularly stable one at lower temperatures (below 500°C) because of the special electronic arrangement which permits its ionization as a donor—equation (1). Further reactions of the SiO_4^+ unit are believed to occur with time. These take place rapidly as the temperature is increased above about 500°C, leading to a loss of the donor. Ultimately a new $(SiO_2)_x$ phase is produced. The kinetics of this process have been discussed in Ref. (6).

When in addition to oxygen, acceptor elements, in particular B, Al or Ga, are present, the reactions are considerably modified.[7] The phenomena occurring, although still not completely understood, are of considerable interest, not only because they bear a close relationship to the reactions which occur when oxygen is present alone, but also because they offer the opportunity to investigate specific effects of the acceptor elements themselves.

This paper presents the results of conductivity and Hall measurements on acceptor-doped silicon crystals homogeneously reacted with oxygen. The results show that a variety of structures form when silicon crystals containing both oxygen and acceptor elements are heated. They also show that electrical methods can be of considerable value in helping to determine the nature of the structures and the reactions which take place. No unique model or process is able to explain quantitatively all of the results. However, it is clear that the reactions which occur in silicon containing oxygen also take place in acceptor-doped silicon containing oxygen, but that, in addition, specific reactions

involving the acceptors occur. A consideration of the possible structures of the products formed and of the possible mechanisms of the reactions, it is believed, will be of use in forecasting the behavior of oxygen as a solute in other crystals and in helping to elucidate the elementary steps in solid state reactions.

2. PROCEDURE AND MATERIALS

Changes in conductivity were measured both at the temperatures of reaction (430–650°C) and at room temperature (25°C). The former procedure was possible only with specimens which remained extrinsic, i.e. were sufficiently highly doped at the temperature of reaction. Such "*in situ*" runs in many cases were measured by means of a recording potentiometer, a method which is convenient especially for kinetics investigations. Conductivity was measured both by means of Al wires welded to the specimens and by means of a two-point probe device. Since the measurements by the two methods agreed within about 1 per cent, no distinction will be made in reporting the results. Most reactions were carried out in air since tests in various gases showed no effects of ambient.

The specimens employed were taken from the central portions of slices cut from pulled (rotated) silicon crystals. The crystals were doped with B, Al and Ga to give acceptor concentrations ranging from about 10^{16} to 10^{18} cm^{-3}. Estimates of oxygen concentration were made by means of infrared absorption for specimens of low acceptor dopings and, in the more heavily doped crystals, by measuring the undoped portions of the crystals at equal diameters. In a number of runs the same specimens were used over again after heating to 1375°C for 5 min and quenching. Except for a small effect on the initial rate of reaction good duplication of the results was obtained. In fact, it was early found desirable completely to remove "growth history" in original specimens before a run by heating as above and quenching to below red heat in 3 sec. This procedure was followed in this paper.

Because of the fact that the acceptors in silicon have fairly deep energy levels and because, during reaction, extra scattering centers are formed, it is necessary to pay particular attention to the interpretation of changes in carrier concentration at room temperature based on conductivity measurements.

To determine carrier concentration in acceptor- and donor-doped specimens it is necessary to have drift mobility values for holes and electrons as a function of concentration. The hole mobilities employed here are based on conductivity measurements on crystals whose Ga contents were obtained by means of gamma-ray counting. Similarly, the electron mobilities were determined from conductivity measurements (25°C) on As-doped crystals in which the As was determined by gamma-ray counting. The data relating carrier concentrations to resistivities obtained in this manner have been combined with already published[8] results to give the curves in Fig. 1.

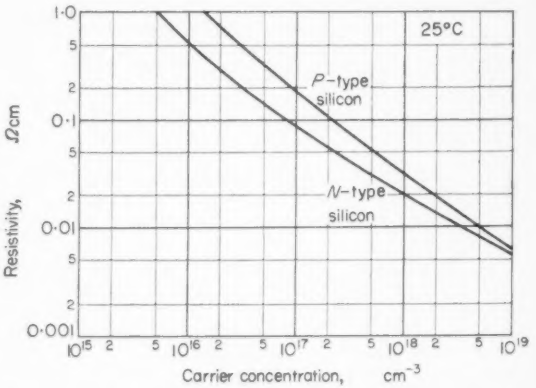


FIG. 1. Relations between carrier concentration and resistivity for silicon containing singly-charged donors or acceptors. The results are based on radiochemical determinations and previously published values[8] of electron and hole mobilities. (25°C.)

The determination of the carrier concentrations in the reacted acceptor-doped specimens presents a somewhat more complicated problem since the SiO_4 units formed may pair or react with the acceptor or remain separate. If the reacted specimens convert to *n*-type then there is good evidence (presented later) that the acceptor scattering is absent and only singly-charged donors are present when the acceptor is fully reacted. Consequently, in this case the electron content is given by the curve in Fig. 1. However, if the reacted specimen is *p*-type, or if there are unreacted acceptors in *n*-type, the resistivity does not give the correct carrier concentration (unless all the donors and acceptors are paired). Hall mobility measurements show that during reaction the hole mobility usually decreases showing that additional scattering centers are introduced. This is always true of specimens that remain strongly *p*-type. In order to make allowances for this, the procedure was adopted of employing for the reacted specimens the mobility corresponding to N_A+N_D. Although this procedure, which assumes the formation of unpaired scatterers is incorrect near intrinsic and in the *n*-type range, the error in these cases is small inasmuch as the concentrations in these ranges are themselves relatively small because of the high resistivities of the reacted specimens.

It remains to determine from the change in carrier concentration what corresponding change occurs in

VOL. 13 1960

$N_A - N_D$, where N_A and N_D are the total acceptor and donor concentrations. First it is necessary, however, to determine N_A, the acceptor concentration in the original *unreacted* specimen, from its hole concentration. This is done using the method of SHOCKLEY[9]. The equation, $p + p_A = N_A$, in which $p_A = f_A N_A$, where f_A is the Fermi function and p_A is the concentration of unionized acceptor, is employed together with equations (2) and (4) below. For electrons equations (3) and (5) below are employed with the equation $n + n_D = N_D$, where n_D is the concentration of unionized donors.

$$p = g_p^{-1} N_V \left(\frac{m_p^*}{m}\right)^{3/2} \exp\left(\frac{E_V - E_F}{kT}\right) \quad (2)$$

$$n = g_n^{-1} N_c \left(\frac{m_n^*}{m}\right)^{3/2} \exp\left(\frac{E_F - E_c}{kT}\right) \quad (3)$$

$$f_A = \frac{1}{1 + g_p^{-1} \exp\left(\frac{E_F - E_A}{kT}\right)} \quad (4)$$

$$f_D = \frac{1}{1 + g_n^{-1} \exp\left(\frac{E_D - E_F}{kT}\right)} \quad (5)$$

In equations (2)–(5) the following values[10] for the quantities shown were employed: $g_p^{-1} = 1/4$, $g_n^{-1} = 1/2$, $N_v = N_c = 2{\cdot}53 \times 10^{19}$ ($T = 300°\text{K}$), $-E_v = E_c = 0{\cdot}545$ (energy measured from middle of band gap),

$$\left(\frac{m_p^*}{m_p}\right)^{3/2} = 0{\cdot}458, \quad \left(\frac{m_n^*}{m_n}\right)^{3/2} = 1{\cdot}13$$

E_A is 0·045, 0·065, 0·072 eV for B, Al and Ga, respectively.[11] E_D for As is 0·049 eV, f_A is the fraction of holes bound to acceptors and f_D is the fraction of electrons on donors.

Suitable allowance must be made for change in E_A and E_D with concentration. In most cases, such corrections are small. Values based on Hall measurements were used in the present work. When both donors and acceptors are present the problem must be solved graphically to obtain E_F and thus f. However, it has been found that in this case about 10 per cent maximum error is introduced by employing the same procedure as described above for N_A but using $N_A - N_D$ instead. In this way values of N_D may be obtained under the assumption that N_A remains constant.

A similar procedure to the above was followed for n-type specimens using equations (3) and (5) to determine the excess N_D. When for either holes or electrons the Fermi level approaches within about 2 kT of the band edge suitable correction should be made following Ref. (9), p. 242. However, this has not been necessary in the present experiments.

Fig. 2 shows a plot of N_A and N_D against carrier concentration calculated as described above. From it the values of N_A or N_D for acceptors or donors of known ionization energy can be obtained from determinations of the carrier concentrations.

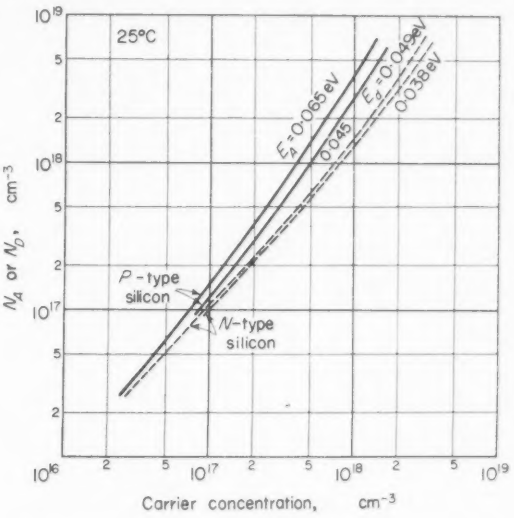


FIG. 2. Plots of N_D or N_A vs. carrier concentration calculated for different donor and acceptor ionization energies. (25°C.)

Measurements of Hall effect were made on bridge-type specimens. These have already been described.[10] Calculations of N_A and N_D were made in the usual manner using the high temperature and low temperature approximations and taking into consideration the effect of the excited states of the acceptors.[12] In calculating carrier concentration from the Hall coefficient, a ratio of Hall to drift mobility of 0·9 was used for holes and 1·2 for electrons. These values gave results in good agreement with the values from conductivity as described above.

3. RESULTS OF CONDUCTIVITY AND HALL EFFECT MEASUREMENTS

(a) *Boron-doped crystals*

Fig. 3 shows the decrease in $N_A - N_D$ as determined from conductivity (25°C) as a function of time of reaction for two B-doped specimens, one containing $8{\cdot}3 \times 10^{16}$ cm^{-3} B and one containing $1{\cdot}1 \times 10^{18}$ cm^{-3} B. In both crystals the oxygen

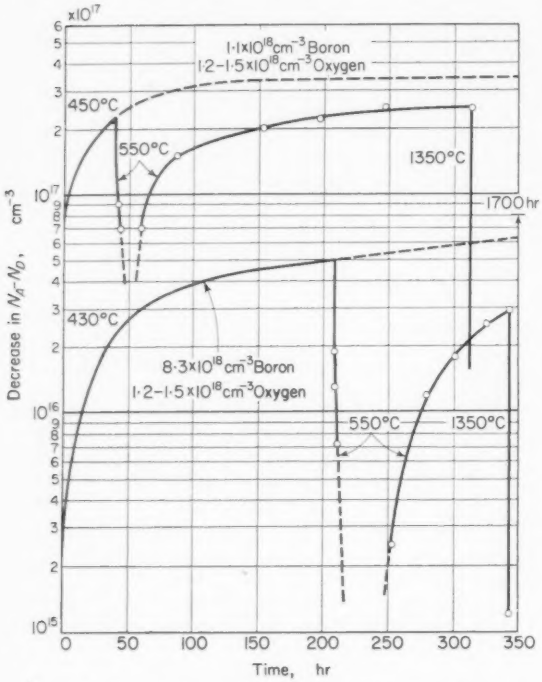


FIG. 3. Reactions at different temperatures for two B-doped silicon crystals containing $1{\cdot}2$–$1{\cdot}5\times10^{18}$ oxygen and $1{\cdot}1\times10^{18}$ cm^{-3} B (upper curve) and $8{\cdot}3\times10^{16}$ cm^{-3} B (lower curve). The decrease in (N_A-N_D), calculated from resistivity measurements (25°C) with the aid of Figs. 1 and 2, is plotted as ordinate. Time in hr at the temperatures indicated on the figure is the horizontal axis.

concentration was between $1{\cdot}2$ and $1{\cdot}5\times10^{18}$ cm^{-3}. The left-hand parts of the curves are for reactions at 430°C and 450°C respectively,* and the right-hand parts are for reactions at 550°C as indicated. A final heat at 1350°C for 5 min returned the specimens to their original states. The following points are to be noted: (i) The initial decrease in (N_A-N_D) is much larger than that (1–3×10^{16} cm^{-3}) observed for undoped silicon containing similar concentrations of oxygen. (ii) The decrease in (N_A-N_D) in the specimen containing $8{\cdot}3\times10^{16}$ cm^{-3} B approaches that of the B-doping itself, whereas that for more concentrated specimens approaches a value of about $3{\cdot}5\times10^{17}$ cm^{-3}. Other runs, not reported here, show that the latter value is essentially the same for higher B concentrations suggesting that this limit is set by the oxygen concentration in the crystals. Further discussion of this effect is given below. (iii) An abrupt increase in holes occurs upon raising the temperature to 550°C in both specimens, followed by less rapid decreases in hole concentration than originally observed at the lower temperatures. The final limits approached by the curves at 550°C fall close to, but somewhat below, those at the lower temperatures suggesting that essentially the same factors (boron and oxygen concentration) are also imposing these limits, but that a different, more stable structure is formed.

* The left-hand curves of Fig. 3 represent average runs on the respective crystals.

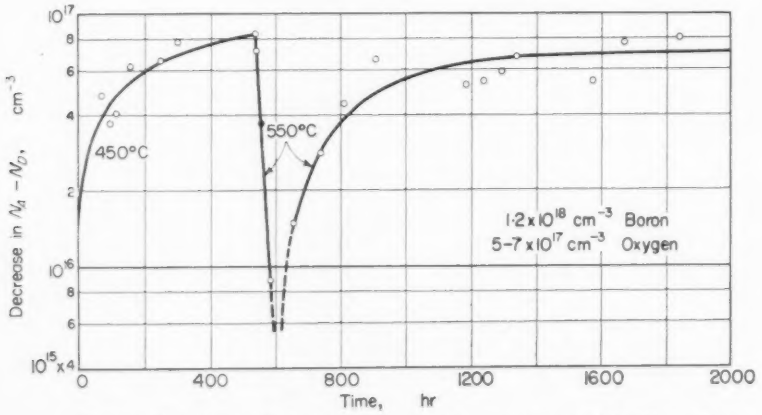


FIG. 4. Reaction of B-doped silicon crystal containing $1{\cdot}2\times10^{18}$ cm^{-3} B and 5–7×10^{17} cm^{-3} oxygen at 450°C and 550°C as determined from resistivity change (25°C). Decrease in N_A-N_D is plotted versus time in hr.

Fig. 4 shows results similar to those of Fig. 3 except that the oxygen content is reduced to between $5–7\times10^{17}$ cm^{-3}. The boron concentration is $1{\cdot}2\times10^{18}$ cm^{-3}. The effect of the lower oxygen concentration, in reducing both the rate and extent of reaction is evident (compare Fig. 3). The same behavior is observed as before, however, with the above, assuming isolated singly-charged donors. This result shows that most of the donors formed are unpaired under these conditions.

(b) *Aluminum-doped crystals—conductivity measurements*

Fig. 5 shows the changes in (N_A-N_D) caused by

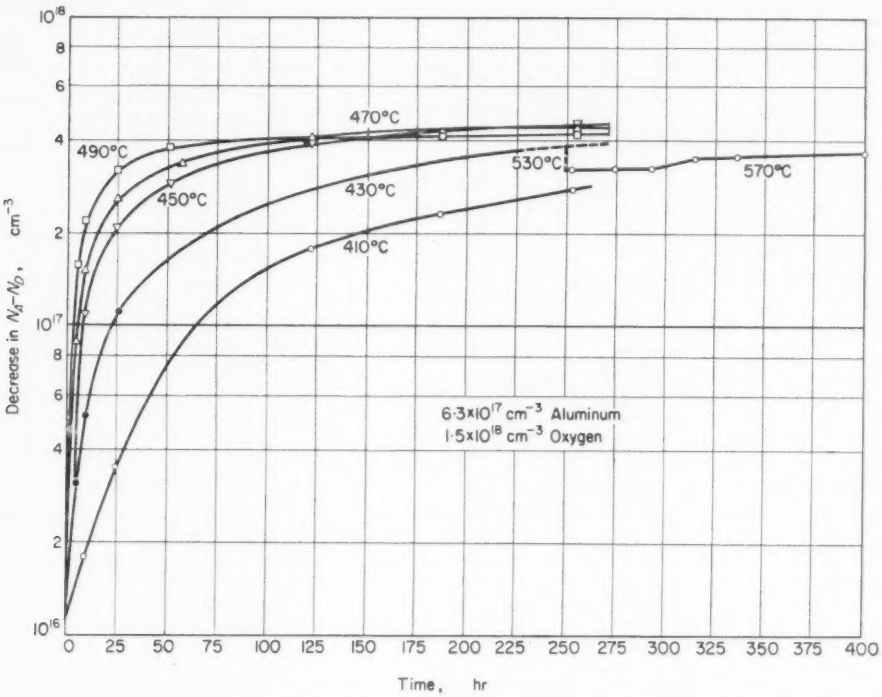


FIG. 5. Reactions at various temperatures of an Al-doped silicon crystal containing $6{\cdot}3\times10^{17}$ cm^{-3} Al and $1{\cdot}2–1{\cdot}5\times10^{18}$ cm^{-3} oxygen based on resistivity measurements (25°C). The effect of heating to 530°C and 570°C is to be contrasted with the behavior of B-doped crystals Figs. 3 and 4. Decrease in (N_A-N_D) is plotted against time of reaction in hr.

namely a rapid increase in holes at 550°C, followed by a slow decrease. The right-hand curve, except for a lower rate, follows the course of the initial curve. Further discussion of these results is given later.

Hall effect measurements were made on a boron-doped ($7{\cdot}8\times10^{16}$ cm^{-3}) crystal after reaction for 90 hr at 430°C. Analysis of the results showed that $5{\cdot}0\times10^{15}$ cm^{-3} acceptors (N_A) were lost and $2{\cdot}1\times10^{16}$ cm^{-3} donors (N_D) were gained. The observed mobility is in satisfactory agreement reacting specimens of Al-doped silicon of 0·074 ohm cm resistivity (25°C) corresponding to $6{\cdot}3\times10^{17}$ cm^{-3} Al and containing $1{\cdot}5\times10^{18}$ cm^{-3} oxygen. Runs were made at a series of temperatures as shown in Fig. 5. The curves all approach a common limit of hole concentration corresponding to a change in (N_A-N_D) between 4·0 and $4{\cdot}5\times10^{17}$ cm^{-3}. This behavior is in contrast to that of B-doped crystals, in results not included here, which show maxima at the higher temperatures (470 and 490°C) similar to those shown by

VOL. 13 1960

undoped crystals.[5] Fig. 5 also shows the effect of an increase in temperature from 430°C to 530°C and later to 570°C. A small, but abrupt increase in hole concentration followed by a slow decrease is noted. This is suggestive of the behavior of B-doped specimens already described in Figs. 3 and 4. However, the small magnitude of the effect in Al as compared to B is significant and will be referred to later.

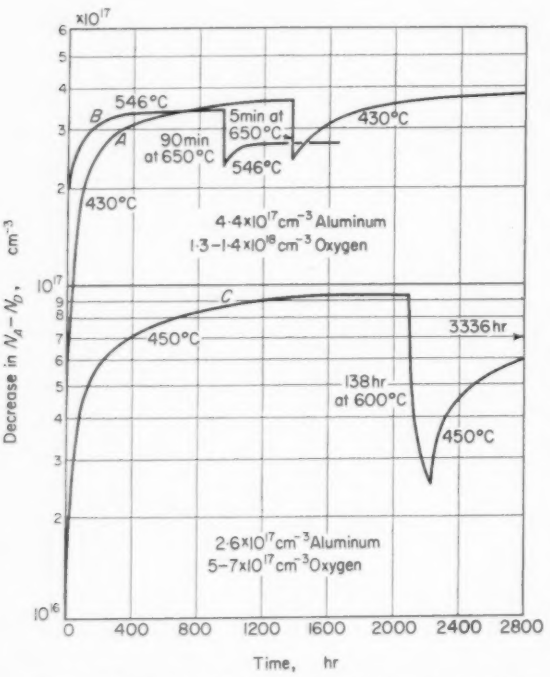


FIG. 6. Effect of 650°C heats on initial reaction products for Al-doped crystals. (a) Initial reaction and final reaction at 430°C, 5 min heat at 650°C. (b) Initial and final reaction at 546°C, 90 min heat at 650°C. (c) Initial and final reaction at 450°C, 138 hr heat at 650°C, low oxygen content crystal ($5–7\times10^{17}$ cm^{-3}). Decrease in N_A-N_D is plotted against time in hr.

The effect of increased temperature upon Al-doped specimens, which have already been reacted to saturation or near saturation at lower temperatures, has been examined in some detail. Fig. 6 shows plots of three different runs.* One run (curve *A*), made at 430°C, was interupted by a 5 min heat at 650°C and continued at 430°C. It is seen that good reversibility is achieved under these conditions. The run at 546°C (curve *B*) was made in the same way, except for the higher temperature and the longer time at 650°C (90 min). Here only a fraction (30 per cent) of the holes are "recaptured". For the third run (curve *C*), which was on a crystal containing $5–7\times10^{17}$ cm^{-3} oxygen, a behavior similar to the first was found. In this case the high temperature heating was at 600°C and for a much longer time (138 hr). As in *B* the extent of the subsequent reaction at 450°C is somewhat less than that of the initial reaction suggesting a loss of some of the reacting components during the long heat at 600°C.

Some added features of these reactions are given in Table 1. In this table are shown the effects of 10 min heats at 550°C, 600°C, and 650°C on specimens which have been reacted at lower temperatures. The changes in (N_A-N_D) are assumed to be due to changes in N_D alone. Several other crystals have been included for comparison. The left-hand columns give the conditions under which the donors were initially formed, the three right-hand columns show the fraction of the initial donors lost at higher temperatures. The following points are to be noted: (i) Compared to B-doped and undoped crystals the Al-doped crystals show much smaller donor loss at 550°C (rows 1 to 4). (ii) At 600°C the donor loss for Al-doped crystals approaches that observed for B-doped and for the undoped crystals provided the formation has been carried out at relatively low temperatures or for relatively short times (rows 3 to 6). If this is not the case, and in particular if the Al-doping is low, then there is very little effect of the 600°C or of the 650°C heat (rows 7 and 8). (iii) Somewhat the same statements as made in (ii) hold for the specimens heated at 650°C. Thus a specimen reacted a short time at 546°C (rows 9 and 10) shows a larger change than one reacted a long time at 546°C (row 11). Also, very long reaction even at a low temperature noticeably reduces the donor loss (row 12). (iv) Lightly doped Al-doped crystals even when reacted for long times, show a much greater donor loss if the oxygen content in the crystal is reduced (row 13). (v) The donors produced in the second reaction in B-doped crystals show considerable heat stability (row 14). (vi) Ga-doped crystals

* Because of the closeness of the experimental points to the curves, no point designations are shown on Fig. 6.

Table 1. Fraction of initial donors recovered at different temperatures ($1{\cdot}2$–$1{\cdot}4\times10^{18}$ cm^{-3} *oxygen*)

Orig. acceptor conc. (cm^{-3})	Anneal T(°C)	Anneal t(hr)	Resulting donor conc. (cm^{-3})	Fraction donors lost* in 10 min at		
				550°C	600°C	650°C
(1) $2{\cdot}4\times10^{17}$ Al	450	282	$2{\cdot}4\times10^{17}$ *i*	0·03		
(2) $1{\cdot}7\times10^{17}$ Al	450	64	$1{\cdot}5\times10^{17}$ *p*	0·03		
(3) $1{\cdot}0\times10^{17}$ B	430	200	$0{\cdot}6\times10^{17}$ *p*	0·30	0·75	0·96
(4) none	430	200	$0{\cdot}2\times10^{17}$ *n*	0·30	0·85	0·99
(5) $4{\cdot}3\times10^{17}$ Al	430	618	$4{\cdot}2\times10^{17}$ *p*		0·52	
(6) $7{\cdot}2\times10^{17}$ Al	430	64	$1{\cdot}8\times10^{17}$ *p*		0·50	
(7) $2{\cdot}8\times10^{17}$ Al	470	755	$4{\cdot}0\times10^{17}$ *n*		0·01	0·01
(8) $1{\cdot}6\times10^{17}$ Al	470	701	$2{\cdot}4\times10^{17}$ *n*		0·00	0·01
(9) $4{\cdot}4\times10^{17}$ Al	546	2	$1{\cdot}1\times10^{17}$ *p*			0·62
(10) $7{\cdot}1\times10^{17}$ Al	546	2	$1{\cdot}5\times10^{17}$ *p*			0·62
(11) $4{\cdot}4\times10^{17}$ Al	546	958	$3{\cdot}4\times10^{17}$ *p*			0·03
(12) $4{\cdot}5\times10^{17}$ Al	430	1380	$3{\cdot}7\times10^{17}$ *p*			0·34
(13) $2{\cdot}6\times10^{17}$ Al[a]	450	2100	$1{\cdot}0\times10^{17}$ *p*		0·29	
(14) $11{\cdot}0\times10^{17}$ B[b]	550	273	$2{\cdot}3\times10^{17}$ *p*			0·06
(15) $3{\cdot}8\times10^{17}$ Ga[c]	430	155	$1{\cdot}4\times10^{17}$ *p*			0·97

* Actually changes in $(N_A-N)_D$ are given. N_A is assumed to remain constant.
(a) Low oxygen content 5–7×10^{17} cm^{-3} oxygen. (b) B-doped crystal after formation of "second donors".
(c) Ga-doped crystal $\sim 1{\cdot}0\times10^{18}$ cm^{-3} oxygen.

Table 2. Comparison of original Al *Concentration before anneal with final donor concentration after anneal* ($1{\cdot}2$–$1{\cdot}5\times10^{18}$ cm^{-3} *oxygen*)

Crystal	Reaction T(°C)	Reaction t(hr)	Original ρ(Ω cm) p-type	Final ρ(Ω cm) n-type	Original $N_A\times10^{16}$ (cm^{-3})	Final $(N_D-N_A)\times10^{16}$ (cm^{-3})
IX 358	500	623	1·14	0·54	1·3	1·0
	,,	,,	0·95	0·28	1·6	2·1
	,,	,,	0·58	0·18	3·2	3·6
	,,	,,	0·42	0·14	4·8	4·9
	,,	,,	0·30	0·11	7·6	6·9
IX 287	490	254	0·49	0·14	4·4*	5·8
	,,	,,	0·24	0·08	11·2*	11·9
	,,	,,	0·18	0·07	17·0*	13·5
	,,	,,	0·17	0·07	18·3*	13·5
	470	755	0·23	0·076	11·1*	11·2
IX 325	600	132	0·15	0·15	20·5	4·7
	,,	,,	0·15	0·13	20·5	5·6
	500	764	0·19	0·10	16·0	8·5
	,,	,,	0·17	0·11	18·5	6·7
	,,	,,	0·15	0·19	21·5	3·3
	,,	,,	0·14	0·28	24·5	2·1
	,,	,,	0·13	int.	27·0	0
	600	10	2·68	1·03	0·5	0·5
VI 1468	600	96	0·135	0·11	25·0	6·7
	700	25	0·134	0·19	25·0	3·3
	900	19	0·134	1·98	25·0	0·3

* Corrected for 5×10^{15} cm^{-3} As compensation.

VOL. 13 1960

behave much like B-doped crystals (rows 3 and 15).

The data of Table 1, as well as other results not included, indicate that the extent of the initial reaction is much more critical for Al-doped crystals in determining their subsequent behavior at 550–650°C than it is for B- and Ga-doped crystals. This indicates that at least two different structures of different thermal stabilities are formed in the Al case, whereas in the B- and Ga-doped specimens only the less stable structures are produced initially. More stable structures, however, form in the second reactions in the latter cases.

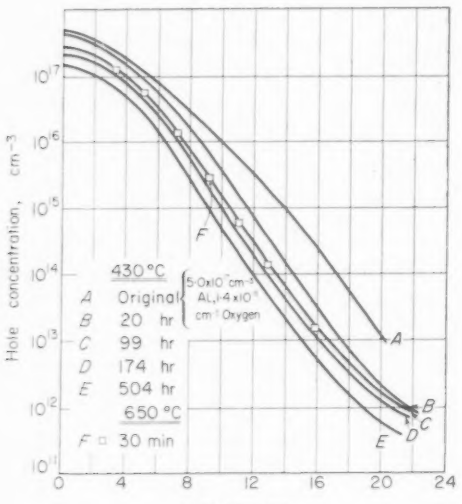


FIG. 7. Hall effect determinations of hole concentration plotted against reciprocal of absolute temperature for an Al-doped crystal containing $5{\cdot}0\times10^{17}$ cm^{-3} Al and $1{\cdot}4\times10^{18}$ cm^{-3} oxygen. Successive curves represent stages in reaction at 430°C. Curve *F* was obtained when bridge was heated 30 min at 650°C after the 430°C reaction.

Table 2 is a compilation of results for a number of Al-doped crystals of different Al contents. In all cases reaction was to completion at the temperatures indicated. It is seen that for Al-dopings up to about 3×10^{17} cm^{-3}, the final specimens all had converted to *n*-type. Several features of the data in Table 2 are noteworthy: (i) Up to a doping of $1{\cdot}7\times10^{17}$ cm^{-3} the number of excess donors per cm^3 is essentially equal to the original Al doping. (ii) Beyond about $1{\cdot}5\times10^{17}$ cm^{-3} the final N_D-N_A is less than the original N_A and an increasing degree of compensation of the Al acceptors occurs. (iii) At about $2{\cdot}7\times10^{17}$ cm^{-3}, compensation for the crystal IX 325 is complete. (iv) In the range of concentration from 1·5 to 3·0 10^{17} cm^{-3} the sum of the original acceptor concentration and the final donor concentration is a constant about equal to 3×10^{17} cm^{-3} for most of the crystals examined.

This follows from the fact that if N_D^E is the excess donor concentration, $N_D^E=N_D^O-(N_A-N_D^O)$, since N_D^O, the number of donors formed by reaction of all of the oxygen present, is compensated by $N_A-N_D^O$ unreacted Al acceptors per cm^3, N_A being the original Al concentration. Thus, the donors due to the oxygen are given by $(N_D^E+N_A)/2$. This concentration, from Table 2 ranges from about $1{\cdot}4\times10^{17}$ cm^{-3} for crystal IX 325 to $1{\cdot}6\times10^{17}$ cm^{-3} for crystals IX 287 and VI 1468. Since 4 oxygens are contained in the donor (SiO_4), this would suggest that at least two SiO_4 units are associated with each Al, inasmuch as the oxygen contents of these crystals are estimated to be between $1{\cdot}2$ and $1{\cdot}5\times10^{18}$ cm^{-3}.* Actually nine to ten oxygen atoms per donor would satisfy the above figures more closely, but the accuracy in the oxygen determinations ($\pm1\times10^{17}$ cm^{-3}) does not justify this close an interpretation of the results.

Increasing the Al-doping above $3{\cdot}0\times10^{17}$ cm^{-3} leaves the specimens *p*-type as has been shown in previous examples (Fig. 5 and 6).

As in the case of B-doped crystals, Al-doped crystals which have lower oxygen contents (5–7×10^{17} cm^{-3}) exhibit reduced reaction rates (curve *C*, Fig. 6). Likewise, the maximum change in carrier concentration is reduced. The latter appears to be somewhat less than proportional to the oxygen content of the crystal although too much confidence cannot be placed in the oxygen content of this crystal.

* *Note added in proof:* Recent measurements of infrared absorption by H. J. Hrostowsky indicate that Al-doped specimens which are reacted to compensation still retain one-half of their original oxygen. If this is confirmed, it indicates that only one SiO_4 unit is associated with two Al atoms at this stage. We are unable at this time to fit this result into the model proposed here.

VOL. 13 1960

(c) *Aluminum-doped crystals—Hall measurement*

Measurements of Hall effect have been made on Al-doped crystals of different concentrations during the course of reaction at 430°C. The most precise results were obtained on a crystal containing $1{\cdot}4\times10^{18}$ cm^{-3} oxygen and $5{\cdot}0\times10^{17}$ cm^{-3} Al (0·11 Ω cm at 25°C). After 500 hr reaction, the bridge specimen was heated to 650°C for 30 min and remeasured to gain additional information on stability.

The results, plotted as p against $1/T$, are shown in Fig. 7. Fig. 8 gives a plot of the results of the calculation of N_A and N_D from Fig. 7, as a function of time of reaction.

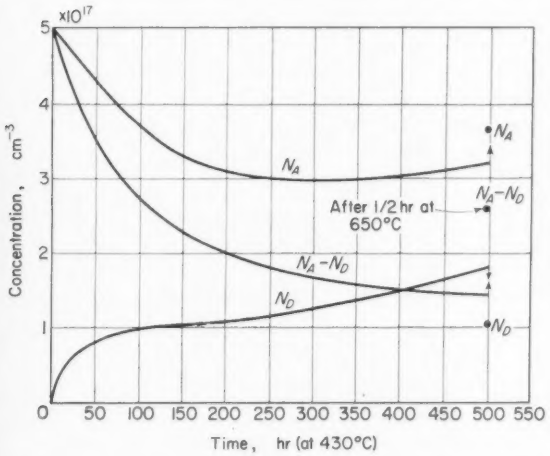


FIG. 8. Plot of N_A, N_D and (N_A-N_D) calculated from curves of Fig. 7 against reaction time at 430°C.

Other Hall effect measurements were made on specimens of Al-doped ($3{\cdot}4\times10^{17}$ cm^{-3}) silicon containing $1{\cdot}2\times10^{18}$ cm^{-3} oxygen reacted at much higher temperatures. Fig. 9 shows these results. Two nearly identical specimens P_1 and P_2, are shown. These were heat-quenched from 1350°C and provided slopes of 0·060 eV corresponding to Al. One specimen, heated for 7 hr at 800°C, turned n-type (N_1) giving $N_D-N_A = 8\times10^{15}$ cm^{-3} and an ionization energy of 0·038 eV. The other specimen, reacted at 695°C for 88 hr, also turned n-type (N_2). For this N_D-N_A was $3{\cdot}2\times10^{16}$ cm^{-3} and the energy was 0·019 eV, precisely one-half that shown by the first specimen. The final curve P_3 was obtained from specimen N_2 after heating 55 min at 1255°C. The specimen turned p-type and gave an activation energy of 0·060 eV corresponding to excess Al. Calculations of N_D and N_A gave $N_D = 3{\cdot}4\times10^{16}$ cm^{-3}, $N_A<10^{10}$ cm^{-3} for N_2 and $N_D = 1{\cdot}88\times10^{16}$ cm^{-3}, $N_A = 1{\cdot}08\times10^{16}$ cm^{-3} for N_1. Since for N_2, N_A is small compared to both n and N_D, one-half the ionization energy for the donor (0·038 eV) is observed (see Ref. (10), p. 30). The fact that different conditions of reaction lead to the same value for

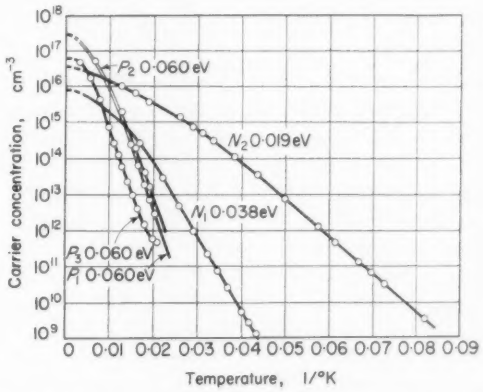


FIG. 9. Hall effect results obtained on two Al-doped specimens P_1 and P_2 containing $3{\cdot}4\times10^{17}$ cm^{-3} Al and $1{\cdot}2\times10^{18}$ cm^{-3} oxygen reacted at various temperatures. Curve N_1 was obtained after P_1 was heated 7 hr at 800°C. Curve N_2 was obtained when P_2 was reacted 88 hr at 695°C. Curve P_3 was obtained when specimen of N_2 was heated 55 min at 1255°C. The P and N designations refer to the specimen types. Carrier concentration is plotted against reciprocal of absolute temperature.

the energy of ionization suggests that it corresponds to a particularly stable form of a donor involving oxygen. Although the value of the energy is close to that for Sb, there seems to be little possibility of attributing it to this impurity. It is interesting that the reaction at 695°C results in almost complete loss of the Al states as far as the Hall measurements show. This suggests that almost all of the Al is paired with donors or contained in neutral complexes.

Two other reactions, one on a crystal containing $6{\cdot}5\times10^{16}$ cm^{-3} Al and $1{\cdot}2\times10^{18}$ cm^{-3} oxygen and one on a crystal containing $9{\cdot}1\times10^{17}$ cm^{-3} Al and $1{\cdot}2\times10^{18}$ cm^{-3} oxygen, were investigated. The results for the former showed the same initial

Table 3. Comparison of calculated and observed mobilities ($9{\cdot}1\times10^{17}$ cm^{-3} Al, $1{\cdot}2\times10^{18}$ cm^{-3} *oxygen*

Reaction time at 430°C (hr)	(N_A-N_D) (cm^{-3})	N_D(cm^{-3})	μ_H(obs.) cm^2/V sec (100°K)	μ_H(calc.) cm^2/V sec (100°K)
0	$8{\cdot}7\times10^{17}$	$0{\cdot}36\times10^{17}$	800	1130
16·5	8·1 ,,	1·0 ,,	730	795
32	7·7 ,,	1·4 ,,	620	679
53	7·0 ,,	2·1 ,,	530	510
117	6·5 ,,	2·6 ,,	500	482
260	6·2 ,,	2·9 ,,	440	454
618	6·0 ,,	3·1 ,,	400	425

trends for N_A and N_D as shown in Fig. 8 and are not reproduced here. A comparison of the calculated (see Section 5) with the observed mobilities for the higher doped crystal above showed the following results (Table 3) for the reaction at 430°C. Because of the high doping separate values for N_A and N_D could not be estimated for this crystal. In making the calculations below, N_A is assumed to remain constant during reaction and all changes in N_A-N_D are attributed to changes in N_D.

The agreement between the calculated and observed mobilities shown in Table 3 suggests that in the heavily-doped crystal largely unpaired donors are formed. This is in contrast to the results obtained on the lighter-doped crystal ($5{\cdot}0\times10^{17}$ cm^{-3}) as shown in Fig. 7 where there is evidence of pair formation (see later discussion).

(d) *Gallium-doped crystals*

Fig. 10 shows results for a 0·24 Ω cm Ga-doped silicon crystal ($\sim1{\cdot}0\times10^{18}$ cm^{-3} oxygen) similar to those shown above for B and Al. The general behavior is again the same although the details differ. The behavior of Ga-doped crystals at higher temperatures *after reaction at* 450°C is similar to that of B. However, the "second rise" in the case of Ga is considerably smaller. In view of this similarity to B-doped crystals no further results will be given. No Hall effect measurements were made on Ga-doped specimens.

Additional conductivity results for Ga-, B- and Al-doped crystals were obtained for reactions in the temperature range 700–1000°C. Some of these for Al will be referred to later. If heating at these temperatures is done after the second generation reaction, the effect is to decrease the electron concentration of *n*-type and to increase the hole concentration of *p*-type specimens. This occurs most readily in Ga-doped crystals leading to a large loss of donors at about 700–800°C. The donors in boron also persist to about 700–800°C, whereas those in Al do not disappear until near 1100°C. Heats at 1375°C, as already mentioned, return the

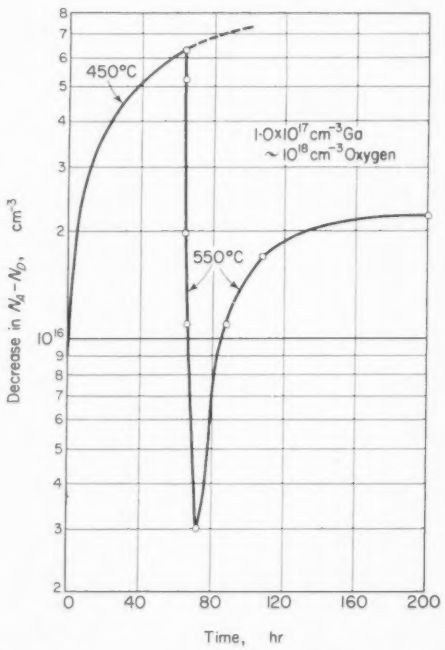


FIG. 10. Decrease in N_A-N_D against time for a Ga-doped silicon crystal containing $3{\cdot}8\times10^{17}$ cm^{-3} Ga and $1{\cdot}0\times10^{18}$ cm^{-3} oxygen. The temperatures are the same as shown in Fig. 3 (upper curve) for boron.

specimens to their original states. It is a significant fact however, that in contrast to Al-doped crystals neither B- nor Ga-doped convert to *n*-type. The behavior of both B- and Ga-doped crystals is erratic above 700° suggesting that perhaps other reactions involving the acceptor take place between 700 and 1200°C.

4. KINETICS OF THE REACTIONS

Attempts have been made to find a kinetic mechanism for the reactions leading to the initial formation of donors in both the boron and aluminum cases. Although simple first and second order kinetics are often able to reproduce the observed curves, evidence from a number of experiments attests to the fact that, similarly to the reactions in undoped silicon containing oxygen, the reactions are complex.[6] Nevertheless, some facts can be learned from a consideration of the approximate kinetics which are helpful in understanding the phenomena taking place.

In general the initial (donor generation) curves conform best to a second order process. This is particularly the case after a heat-quench from 1375°C. A number of kinetics runs in the temperature range 428–550°C were made on the same specimen of Al-doped silicon ($6{\cdot}0\times10^{17}$ cm^{-3} Al, $1{\cdot}4\times10^{18}$ cm^{-3} oxygen) by reversing the reaction at 1375°C between runs. It was found that the rate of quench from 1375°C affected the rate of the first part of the reaction, the rate being accelerated by a faster quench. The effect of temperature on the over-all donor generation reaction, however, was found to be surprisingly small, the apparent activation energy being about 30 kcal calculated either on the basis of initial reaction rate or for second order kinetics. Runs made at the temperature of reaction agreed well with those measured by quenching to room temperature indicating that multiple ionization of the donor, at least in the range up to 500°C, is absent. Thus the donors formed appear to have but a single level.

An investigation has also been made of the kinetics of the reactions causing the abrupt increases in holes (decreases in donors) shown in Fig. 3 for the B-doped crystals. These reactions are first order in the donor concentration. The results, which have been reported elsewhere,[13] are shown compared to the behavior of undoped crystals in Fig. 11. They show that the reduced first order rate constants (plotted as k/T in Fig. 11) are linear against $1/T$ and give a common activation energy of 65 kcal. This strongly indicates that the donors formed in the low temperature range are very similar for the doped and undoped

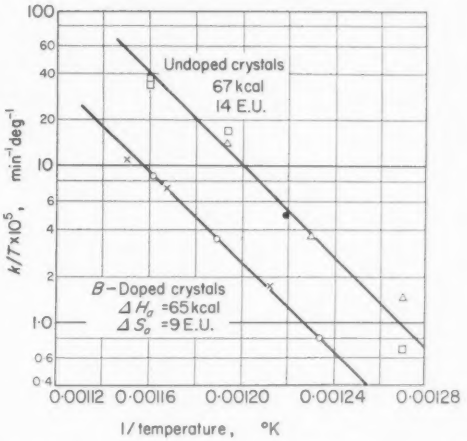


FIG. 11. Plots of reduced first order rate constants for the donor disappearance reactions in undoped and in B-doped ($1{\cdot}1\times10^{18}$ cm^{-3}) silicon crystals as a function of $1/T$. The oxygen content was $1{\cdot}0$–$1{\cdot}2\times10^{18}$ cm^{-3}.

crystals as has been suggested above. The slower rates for the B-doped crystals as compared to the undoped crystals was originally believed to indicate the formation of deep-lying acceptors.[13] However, it now seems more likely that the slower rate may be characteristic of the reaction itself. It also appears likely that the activation energy (65 kcal) is that of the dissociation reaction rather than of the oxygen diffusion.[14]

The results of conductivity measurements, shown in Fig. 6 are interesting in this connection, particularly the comparison of the two runs, curve *B* at 546°C and curve *A* at 430°C, on the same crystal. The former at 650°C shows an order of magnitude greater initial rate of loss of donors than the latter, run at 430°C, which agrees fairly well with Fig. 11 for B-doped crystals. (This difference in the vertical portions of the curves is not visible on the scale used in Fig. 6.) Also, the 546°C run shows relatively small "recovery" compared to the 430°C one, which is essentially reversible. These results again indicate that

VOL. 13 1960

structures of different stabilities are formed in the initial runs at 430 and 546°C. The effect of reduced oxygen content as shown in the run at 450°C of Fig. 6 (curve *C*) is likewise interesting. In spite of a long reaction at 600°C, a large part of the loss is recovered suggesting that the side reactions are less effective in the more dilute solution (compare curve *B*).

5. DISCUSSION OF HALL MOBILITY RESULTS

Plots of the Hall mobility ($R\sigma$) against temperature in °K (Fig. 12) show a large influence

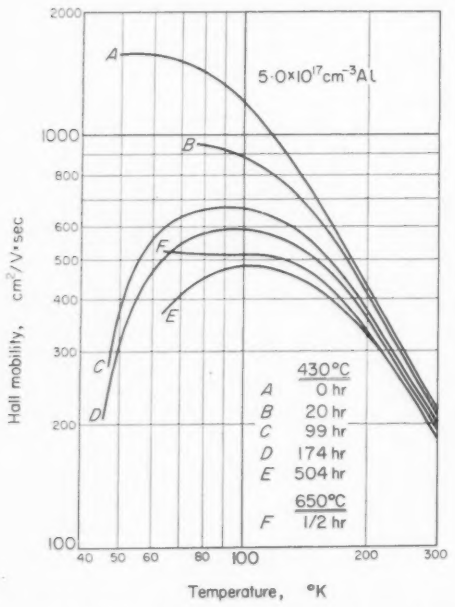


FIG. 12. Plot of change of Hall mobility versus absolute temperature for successive stages of reaction (*A–F*) at 430°C. The curves shown here correspond to those of Fig. 7.

of impurity conductivity at low temperatures. The presence of this effect is in agreement with the formation of compensating or "internally-ionized" solutes of the sort postulated above and might be expected to be accentuated by the large size of the impurity structures.

In estimating the number of scattering centers in *p*-type specimens at low temperatures (100°K), it is assumed that each donor formed ionizes one acceptor and, if the centers are paired, no increase in scattering occurs. However, two new scattering centers are produced if the donors formed are unpaired. Since only the unpaired donors are counted in the Hall effect measurements, only these need to be considered in the calculations of mobility.

Calculations of Hall mobilities have been made using the procedure of SHOCKLEY[9] and CONWELL[15].

$$\mu_1 = 4{\cdot}7\times10^{17}(m_p/m_p^*)^{1/2}\times \times T^{3/2}/N_I \ln(1+4{\cdot}5\times10^8T^2/N_I^{2/3}) \quad (6)$$

where μ_I = the ionized impurity scattering mobility
N_I = concentration of ionized impurities (single-ionized and unpaired)
m_p^* = the effective mass of holes = $0{\cdot}56\, m_p$

Values for μ_T, the lattice scattering mobility have been obtained from Ref. (16).

$$\mu_T = 5{\cdot}67\times10^{17}T^{-2} \quad (7)$$

Results for a crystal containing ($9{\cdot}1\times10^{17}$ cm^{-3} Al) are shown in the last column of Table 3. Except for the unreacted crystal, the calculated mobilities are in good agreement with those observed. The higher value of calculated mobility for the original crystal may be attributable to neutral impurity scattering (oxygen) which will make a small contribution to the total scattering.[16] The changes in mobility in this specimen are therefore in agreement with singly-charged SiO_4^+ donors being formed which for the most part are unpaired.

Similar calculations have been made for the other Hall runs. The results for the run on the Al-doped crystal ($5{\cdot}0\times10^{17}$ Al) shown in Figs. 7, 8 and 12 are particularly interesting. In this case, as shown in Fig. 8, calculations of N_D and N_A show that heating 1/2 hr at 650°C after the run at 430°C produces a decrease in N_D of $1{\cdot}0\times10^{17}$ cm^{-3} and increase in N_A of $7{\cdot}5\times10^{15}$ cm^{-3}. This—as Fig. 7*F* shows—in effect returns the specimen to its previous state corresponding to curve *C*, Fig. 7. However, the mobility—as Fig. 12, curve *F*, shows—does not return to that of curve *C*, and, in addition, a loss of the impurity conduction effect occurs. The decrease in N_D and increase in N_A are in agreement with the reduced compensation and therefore reduced impurity conduction. However, the low value of the mobility observed at

100°K is not in agreement with that calculated from measured N_D and must be therefore attributed to scattering by the new products (neutral) formed in the reaction at 650°C.

Although the mobility results suggest that the donors are largely unpaired when heavily-doped specimens are reacted, this is not true of the crystals containing 5×10^{17} cm^{-3} Al as shown by the decrease in N_A (Fig. 8). Furthermore, the mobilities calculated on the assumption that the N_D donors are unpaired agrees with observations showing that in this case, both paired and unpaired donors form. Measurements have recently been made which show that in *n*-type crystals an increase in mobility actually occurs, thus furnishing further proof that Al^- and SiO_4^+ donor centers react to produce a net decease in scattering.

VOL. 13 1960

6. DISCUSSION OF THE NATURE OF THE REACTIONS

It is obvious from the variety of phenomena taking place during reaction that no one simple process can explain all of the results. It is already known that the reactions occurring in silicon crystals containing oxygen alone, involve successive, and possibly also parallel, partial reactions leading eventually to electrically inactive structures and precipitate phases.[6] There is also considerable evidence that the SiO_4 unit acts as a donor,[4] perhaps as a consequence of the interaction of the nonbonding *p*-orbitals of the oxygens.

(a) *The SiO_4 unit in acceptor-doped crystals*

The fact that the decrease in $(N_A - N_D)$ is equal, within the error of the experiments, to one-fourth of the oxygen content when acceptor-doped crystals are reacted is strong evidence for the formation of the SiO_4 donor. Such behavior is shown for B-doped crystals in Fig. 3 and for Al-doped crystals in Fig. 5 and Table 2 and has been observed in many other examples not reported here. The results for Ga-doped crystals (Fig. 10) are incomplete. Exhaustion of oxygen can only of course occur when sufficient acceptor concentration is present and the reaction is run to completion. The greatly accelerated rate and extent of the formation of the donor in doped as compared to undoped silicon suggests an interaction of the donor electrons with the holes present or of the donor with the acceptor ion itself. Further evidence for the presence of the SiO_4 donor is provided by the similar behavior of specimens containing it, both doped and undoped, when they are subjected to temperatures above 500°C. The rather surprising result that the donor formed with Al contains eight oxygen atoms within the error of the experiments (Table 2) also indicates persistence of the tetrahedral unit.

The above facts, the results on stability and the Hall results, suggest that the reactions occur in stages in which the SiO_4 unit plays a dominant role and that, through the agency of mobile oxygen atoms with which the unit maintains equilibrium, it is able to appear at various sites in the crystal. The first site chosen appears to be similar to that selected in undoped crystals and is not apparently adjacent to the acceptor. This latter site, however, seems to be favored later in the reaction. On the other hand, it seems immediately preferred when the crystal becomes intrinsic or *n*-type.

In the following discussion we shall assume that the SiO_4 unit is able to act alone as a singly-charged donor, that it may also form adjacent to an acceptor and so produce a pair which may not be a donor but may be neutral, and that it may form in duplicate adjacent to an acceptor, in which case it acts as a singly-charged donor (triplet). The formation of a pair is indistinguishable from a separate SiO_4 donor except for the effect on the Hall results. Likewise, the triplet donor may be regarded as two SiO_4 donors associated with the acceptor even though only one ionizable electron is observed. Thus although the combined SiO_4 unit may not have the properties of a donor once it has reacted with an acceptor, it will be convenient when discussing the results to assume that it retains its donor properties even in the above combinations.

(b) *Effect of the hole–electron equilibrium*

Let us examine how far the above assumption is able to go in explaining the experimental results presented above. First, the cause of the large increase in donors shown by acceptor-doped silicon containing oxygen as compared to undoped silicon containing comparable concentrations of oxygen is of interest. Inasmuch as no parallel increase occurs when Group V donor elements are added as doping agents to silicon, one is tempted

to explain the result as caused by the operation of the hole–electron equilibrium on the donor solubility.[17]

$$SiO+O\rightleftarrows SiO_2+O\rightleftarrows SiO_3+O\rightleftarrows SiO_4\rightleftarrows SiO_4^+ + e^-$$

$$B\rightleftarrows B^- + e^+ \qquad e^+ + e^- \rightleftarrows e^+ e^- \qquad (8)$$

where e^+ is the hole, e^- an electron. It was shown in Ref. (17) that for the acceptor to influence the donor solubility its concentration should exceed $\sqrt{K}$, where K is the hole–electron equilibrium constant. This condition is met in many of the experiments here as for example in the upper left-hand curve in Fig. 3 where the B-doping exceeds the $\sqrt{K}$ (4×10^{16} cm^{-3}) at 430°C. However, it is also met in the right-hand part of the curve at 550°C ($\sqrt{K}=2\times10^{17}$ cm^{-3}) where the abrupt decrease in donors occurs. The latter decrease must therefore be the result of reaction at the higher temperature rather than an effect of the hole–electron equilibrium. Furthermore, as the lower curve in Fig. 3 shows, B is able to increase the number of donors even when its concentration falls below the concentration corresponding to $\sqrt{K}$. Thus, in addition to the hole–electron effect, B and the other acceptors appear to also have specific effects, perhaps related to the formation of stable pairs.

$$Si+4O+A\rightleftarrows SiO_4^+A^- \qquad (9)$$

Such pairs, which can only form through the migration of oxygen available from the set of equilibria shown in (8), are also able to increase the donor solubility[18] and could explain the rise in the donor concentration beyond that required by the hole–electron equilibrium alone.

Even in the "low-oxygen" crystals shown in Figs. 4 and 6, the concentration of donors formed is an order of magnitude greater than that produced in the absence of B or Al, although on account of the operation of the equilibria, equation (8), the reduced oxygen concentration is unable to force the reaction as far to the right as in the case of the "high-oxygen" crystals, for example, as shown in Fig. 3, curve *B*. Again, the decrease in donors at 550°C (Fig. 4) cannot be explained as simply as a solubility effect since $\sqrt{K}=2\times10^{17}$ cm^{-3} at this temperature and over $1{\cdot}0\times10^{18}$ cm^{-3} of *B* is present.

(c) *The second "donor rise" in* B- *and* Ga-*doped silicon*

A most interesting effect shown strikingly by the B-doped crystals (Figs. 3 and 4) and, to a lesser extent by the Ga-doped and by the Al-doped crystals, is the abrupt decrease in donors at 550°C followed by a second generation of donors (loss of holes) at the same temperature. This effect is also found in undoped crystals, but without a second rise in donors, although small such increases have been observed.[5]

This behavior indicates that the equilibria (equation 8) are strongly shifted to the left at about 500°C resulting in a freeing of oxygen from the donor and a loss of its donor properties. The ensuing rise in donors (decrease in holes) is more difficult to explain. It is significant that in each case (Figs. 3 and 4) the curves rise to almost their original values. This strongly suggests that the second structure is merely a rearrangement of the first. Such a rearrangement could occur through migration of oxygen from a less stable to a more stable configuration. There are a number of possibilities. For example stable triplets (donors), SiO_4^+ A^- SiO_4^+, could form as has been suggested or neutral compositions containing the acceptor, perhaps AO_4, or more complex compositions involving a lower oxygen to acceptor ratio, may be produced. However, no direct evidence for these specific structures has been obtained.

The behavior of Ga, which has been the least investigated is similar to B in almost all respects so far observed. In the crystal shown in Fig. 10, neither the final donor concentrations nor the oxygen concentration were known, although the latter is estimated to be $\sim10^{18}$ cm^{-3}. Consequently, it is not known whether the reaction will proceed to complete exhaustion of the oxygen present although this appears unlikely. Evidence from infrared that Si–O–Ga units form has been given in Ref. (11).

(d) *Nature of the reactions in* Al-*doped silicon*

It has been mentioned that the reactions exhibited by Al do not appear to differ fundamentally from those of B and Ga. For example, a "second rise" in donors is also observed for Al-doped crystals (Fig. 5) although the magnitude is considerably less than that usually observed for B. This suggests that there are fewer isolated SiO_4

VOL. 13 1960

donors formed in Al-doped than in B-doped crystals under similar conditions of reaction. Why the Al atom should serve as a more favorable site than the B atom is not known. Possibly there is less lattice strain about the Al.

Of special interest is the fact that Al-doped specimens are able to turn *n*-type upon reaction. As indicated previously in this section, this behavior cannot be caused by the hole–electron effect or by pair formation. The assumption of a triplet, $SiO_4^+ Al^- SiO_4^+$, agrees with the donor requirement as well as the requirement that approximately eight oxygens are present for each Al atom. However, a structure of the kind $AlO_4 . SiO_4^+$ would also meet these requirements.

The results of Fig. 6, curves *A* and *B*, and Table 1 show that at least two structures of different thermal stability are formed. The more stable of these may correspond to the donor (triplet) above. The less stable is similar to the structures formed in B-doped and undoped silicon and is therefore presumably a single SiO_4 unit or a pair.

Evidence for the formation of pairs is given by the Hall measurements which likewise indicate the production of individual SiO_4 donors. Thus Fig. 8 shows that N_A decreases by 2×10^{17} cm^{-3} for the first 300 hr of reaction for the specimen described there. Such a decrease in N_A can only be explained by the formation of acceptor–donor pairs which would not be counted in the Hall measurement. Also the immediate rise of N_D at the beginning of the reaction suggests the production of separate donors as does the decrease at 500 hr as a result of heating at 650°C. The increase in N_A, occurring at the same time, suggests a dissociation of some of the donors of the pairs. Finally, the slow rise in N_D beyond 150 hr may indicate that the later donors formed are the more stable triplets. Thus, the Hall results can be qualitatively accounted for by the assumption of three structures based on the SiO_4 unit. Just why the different structures should have different stabilities is not known. The increased stability cannot be attributed to coulombic forces since these are small.[18] It is possible that the stability is due to resonating structures.

The Hall results obtained on crystals of Al-doped silicon reacted at still higher temperatures (Fig. 9) show that long reaction of Al-doped crystals at 700°C or above, results in small levels for the N_A and N_D concentrations and suggests that much of the Al is then contained in outwardly neutral oxygenated compounds when, as in this case, sufficient oxygen to react with the Al is present.

Al-doped specimens which have been reacted to *n*-type not only lose the donor on prolonged heating at 700°C, but slowly turn *p*-type showing the presence of excess Al acceptors (Fig. 9). This result is achieved much more rapidly at higher temperatures (1100°C) and above. Fast reactions occur at these high temperatures which have not been investigated. However, as mentioned all the reaction products appear to decompose rapidly at temperatures near the m.p. of silicon. Thus, the solutes are again dispersed and are apparently incapable of recombining again except through the agency of the SiO_4 unit which is unstable except below about 500°C.

7. OXYGEN DIFFUSION

In the above discussion the diffusivities of the acceptor atoms have been disregarded in comparison with the diffusivity of oxygen. A simple calculation shows that oxygen must have a diffusion constant of 10^{-15} to 10^{-18} cm^2/sec in order to have the required mobility at 400–500°C for the reactions described. This is considerably greater than the calculated diffusion constants for the acceptors which for Al, for example, is 10^{-23} cm^2/sec at 450°C, and suggests that one is justified in regarding the acceptors as fixed at these lower temperatures. The curve for Al-doped silicon at 430°C in Fig. 6 shows that it is possible to isolate partial processes which are reversible and which follow precise first order kinetics. This, as has been mentioned, suggests that oxygen can be dissociated from a donor or a donor–acceptor pair and be recombined. The process is analogous to diffusion of a mobile particle to a sink and, provided the particle is not removed too far, may be expected to be first order.[18] The value of the diffusion constant, *D*, for oxygen derived from the above first order reaction using

$$D = k/4\pi RN \qquad (10)$$

comes out to be $1{\cdot}7\times10^{-17}$ cm^2/sec in good agreement with other estimates.[6] In equation (10) *D* is the diffusion constant, *R* is the diameter of the reacting center (here taken as 5A), *N* the

concentration of the center and k, the specific rate constant. In this case N was 10^{17} cm^{-3} and $k = 1{\cdot}06 \times 10^{-6}$ sec^{-1}.

So far it has not been possible to carry out the reverse reaction shown in Fig. 6 at different temperatures and so to obtain the activation energy for the diffusion of oxygen in order to compare it with the kinetics of the initial donor formation mentioned earlier as well as with the diffusion results at high temperatures.[19]

The mechanism by which oxygen migrates in the crystals appears to involve an interstitial movement in which strong bonds are broken and reformed. It should be possible to obtain further information on this mechanism by means of kinetics measurements on reactions such as shown in Fig. 6.

8. SUMMARY AND CONCLUSIONS

Investigation of the solid state reactions which occur in silicon crystals between oxygen, silicon, and acceptors has shown that specific structures, some of which behave as donors, are formed. The mechanisms and the precise structures have not been determined. However, most of the observations can be explained by the assumptions that the SiO_4^+ unit behaves always as a single-leveled donor and that stable combinations of it with the acceptor elements, in particular, $SiO_5^+ A^-$ and $SiO_4^+ A^- SiO_4^+$ are able to form. Two of the above compounds are covalent molecular donors; the other a neutral pair. None of these, however, is supposed to exist in a true equilibrium state with the others but each represents a metastable product subject to transformation into successively more stable structures through the agency of atomic oxygen with which SiO_4 maintains a true equilibrium.

The above molecular ions, particularly the proposed triplet, are supposed to gain particular stability by virtue of resonating structures. The formation of the SiO_4^+ donor is favored in heavily-doped p-type solutions at low temperatures. Higher temperatures, long reaction times, and low acceptor concentrations appear to favor a pair and triplet. In near intrinsic or n-type solution, however, only the more stable structure forms. Although many other (especially intermediate) structures no doubt exist in the course of the reactions which take place such as, for example, $SiO\ A^- SiO_4^+$, $SiO_2\ A^- SiO_4^+$, etc., these are not detectable by means of electrical measurements and their presence is only inferred from subsidiary observations. It is conceivable also that structures of the type AO_4°, $AO_4^\circ\ SiO_4^+$, may be formed in addition to, or instead of, those mentioned above. These do not, however, explain the observations as well and until more specific structural information is available the true nature of the compounds must remain in doubt.

The three acceptors B, Al, and Ga although reacting according to the same general pattern, exhibit definite differences. Only Al is capable of reacting completely to form the stable donor. B and Ga can be made to form an unstable structure at lower temperatures which slowly transforms to more stable structures at higher temperatures.

A way has been indicated whereby more precise kinetic data may be obtained. Such kinetic studies should prove of great value in helping to determine more definitely the mechanisms of the elementary reactions here discussed.

Acknowledgement—The authors wish to thank PAUL FREELAND and E. D. KOLB who supplied the crystals used in this work, and J. P. MAITA for making the Hall measurements of Fig. 9. The measurements of oxygen content were kindly made by W. KAISER and G. H. WHEATLEY. Particular thanks are due to N. B. HANNAY for useful suggestions.

REFERENCES

1. FULLER C. S., DITZENBERGER J. A., HANNAY N. B. and BUEHLER E., *Phys. Rev.* **96**, 833 (1954); *Acta Met.* **3**, 97 (1955).
2. KAISER W., KECK P. H. and LANGE C. F., *Phys. Rev.* **101**, 1264 (1956).
3. KAISER W. and KECK P. H., *J. Appl. Phys.* **28**, 822 (1957).
4. KAISER W., *Phys. Rev.* **105**, 1751 (1957).
5. FULLER C. S. and LOGAN R. A., *J. Appl. Phys.* **28**, 1427 (1957).
6. KAISER W., FRISCH H. L. and REISS H., *Phys. Rev.* **112**, 1946 (1958).
7. FULLER C. S. and DOLEIDEN F. H., *J. Appl. Phys.* **29**, 1264 (1958).
8. PRINCE M. B., *Phys. Rev.* **93**, 1204 (1954); DEBYE P. P. and KOHANE T., *ibid.* **94**, 724 (1954); *ibid.* **105**, 522 (1957); MORIN F. J. and MAITA J. P., *ibid.* **96**, 28 (1954).
9. SHOCKLEY W., *Electrons and Holes*. Van Nostrand, Princeton, N.J. (1950).
10. HANNAY N. B. (Editor), *Semiconductors*. Reinhold, New York (1959).
11. HROSTOWSKI H. J. and KAISER R. H., *J. Phys. Chem. Solids* **4**, 148, 315 (1958).

12. PUTLEY E. H., *Proc. Phys. Soc. Lond.*, **72**, 917 (1958).
13. FULLER C. S., *Chem. Rev.* **59**, 65 (1959).
14. WAITE T. R., *J. Chem. Phys.* **28**, 103 (1958).
15. CONWELL Esther M., *Proc. Inst. Radio Engrs., N.Y.* **40**, 1327 (1952).
16. LOGAN R. A. and PETERS A. J., *J. Appl. Phys.* **30**, 1627 (1959).
17. REISS H. and FULLER C. S., *J. Met. N.Y.* **8**, 276 (1956).
18. REISS H., FULLER C. S. and MORIN F. J., *Bell Syst. Tech. J.* **35**, 535 (1956).
19. LOGAN R. A. and PETERS A. J., *J. Appl. Phys.* **28**, 819, 1419 (1957).

VOL.
13
1960

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 13. pp. 204–211. Printed in Great Britain.

ON THE ISOTOPIC EFFECTS IN THE FERROELECTRIC BEHAVIOUR OF CRYSTALS WITH SHORT HYDROGEN BONDS

R. BLINC

J. Stefan Institute, Ljubljana, Yugoslavia

(*Received* 24 *July* 1959)

Abstract—An attempt to correlate the isotopic effects in the ferroelectric properties of hydrogen-bonded crystals with the observed anomalies in their i.r. and n.m.r. spectra has led to the conclusion that the protons in the hydrogen bonds are tunnelling in double-minimum potential fields. The ferroelectric transition is assumed to be the result of a deformation of the protonic distribution due to electrostatic interactions. It has been found that a quantum-mechanical extension of MASON's and DEVONSHIRE's long-range-forces model is able to explain, by the same mechanism, the results of i.r. and n.m.r. spectroscopy as well as the dependence of the Curie point on the mass of the hydrogen isotope, the sharper increase of the spontaneous polarization with falling temperature and the larger value of the spontaneous polarization at absolute zero for deuterated than for hydrogen compounds. The proposed model predicts that a ferroelectric transition occurs only if the dipole–dipole interactions, expressed as frequencies, are greater than the tunnelling frequency on the lowest vibrational level.

1. INTRODUCTION

THE ferroelectric properties of KH_2PO_4, KH_2AsO_4 and isomorphous ferro- and antiferroelectric compounds with short hydrogen bonds exhibit a much discussed, but as yet unexplained, isotopic effect. In the case of KH_2PO_4, for instance, on deuteration the Curie point is raised by as much as 89°, the increase of the spontaneous polarization with falling temperature becomes much sharper and the value of spontaneous polarization at absolute zero is nearly doubled. Both SLATER's[1] and the DEVONSHIRE–MASON treatments[2, 3] of the nature of ferroelectricity in KH_2PO_4, which are based on double-minimum protonic potential functions, are essentially classical. Their theories are therefore unable to explain the observed isotopic effects. The same argument holds also for the work of GRINDLEY and TER HAAR[4], which is an extension of SLATER's theory.

PIRENNE[5] treated the problem quantum-mechanically, but used a single-minimum potential field. He tried to explain the isotopic effect by assuming that the protons are free to move in a one-dimensional square well. As the vibrational levels are lower for deuterium than for hydrogen, the Curie temperatures are higher for the deuterated compounds. But the resulting protonic vibrational levels do not agree with those derived from i.r. measurements,[6] these being approximately ten times higher. Moreover, a square-well model does not explain the appearance of two OH stretching bands instead of one. Therefore this approach can not be considered satisfactory.

Recently REID[7] in his semi-empirical treatment of the hydrogen bond put forward another atomic model for the ferroelectric effect in hydrogen-bonded crystals, but his model too is at variance with the observed i.r. and n.m.r. data.

It is known that the hydrogen motions are fundamental to the ferroelectric or antiferroelectric phenomena in these crystals, and that the spectral consequences of potential fields with a single minimum or two minima are different. Thus the spectra should help to determine the form of the potential function governing the motion of the

VOL. 13 1960

protons and so elucidate the nature of the isotopic effects. However, no attempt has as yet been made to relate the isotopic effects in KH_2PO_4 to the results of i.r. and n.m.r. spectroscopy.

Both i.r. and n.m.r. spectra of KH_2PO_4 show striking anomalies. Instead of one there are two ν_{OH} bands in the i.r. as well as in the Raman spectrum. Moreover, these bands do not change appreciably on going from the non-ferroelectric to the ferroelectric phase. This occurs in spite of the apparent change of 0·20 Å in the OH bond lengths,[8] which should cause, according to BADGER's rule,[9] a shift of $\sim 400\ cm^{-1}$ in the position of the ν_{OH} band. The p.m.r. spectra of KH_2PO_4 exhibit a line-width transition[10] which clearly shows the existence of protonic motions. The transition point of the line width does not coincide with the Curie temperature, as would be required by SLATER's or REID's theories. The correlation time is short and nearly temperature-independent at low temperatures.

In order to determine whether these effects are limited to KH_2PO_4 only, or whether they are characteristic of all ferroelectrics or antiferroelectrics with short hydrogen bonds, the i.r. and n.m.r. spectra of KH_2PO_4, KH_2AsO_4, $NH_4H_2PO_4$ $NH_4H_2AsO_4$, $Ag_2H_3IO_6$, $(NH_4)_2H_3IO_6$ and their deuterated analogues have been investigated, as well as those of NaH_2PO_4, $CaHPO_4$ and $BaHPO_4$. In the i.r. spectra of all these compounds two ν_{OH} bands in the 1800–2800 cm^{-1} region, 200–400 cm^{-1} apart, have been found. These bands suffer only small shifts on going through the ferroelectric transition points. In addition the p.m.r. spectra of these compounds show the same characteristic behaviour as KH_2PO_4, and they prove unambiguously the existence of rapid protonic motions. The details of this work have been[6] or will be published elsewhere.

From the foregoing it may be concluded that the above-mentioned spectral "anomalies" are inherent properties of ferroelectrics with short hydrogen bonds. The purpose of the present paper is to relate them to the nature of the ferroelectric transition in these compounds.

However, it must be pointed out that, according to the proposed model, these anomalies are a necessary but not sufficient condition for the occurrence of ferroelectric or antiferroelectric transitions.

2. ATOMIC MODEL

Though it has been pointed out that only a BETHE type theory,[11] including both short-range and long-range forces, could explain in a satisfactory manner the whole complex of ferroelectric phenomena, it seemed worthwhile to investigate the extent to which the simple local field theory, as developed by DEVONSHIRE[3], could account for the isotopic effects and the observed spectral anomalies, if treated quantum mechanically.

The proposed model is based on the assumption that the protons in the hydrogen bonds O–H···O are tunnelling in a double-minimum potential field. The minima are of equal depths in the non-ferroelectric phase and unsymmetrical, owing to dipole–dipole interactions, in the ferroelectric form. The total polarization of the crystal is the sum of the protonic polarization (P_d)* which triggers the transition, and the ground polarization ($\gamma\mathscr{F}$), as revealed by the displacements of P and K ions in the case of KH_2PO_4. Since we are primarily interested in the isotopic effects, little attention is paid to the ground polarization. Therefore it is simply included in the factor γ, which means that it is treated in the same way as the electronic polarizability. Further we assume, following DEVONSHIRE, that there is no short-range correlation between the positions of neighbouring dipoles.

The potential field of a hydrogen-bonded proton is not precisely known, but may perhaps be adequately described by:

$$V = V_1 = \tfrac{1}{2}k(x+l)^2 \qquad \text{for } -\infty \leqslant x \leqslant b$$

and

$$V = V_2 = \tfrac{1}{2}k(x-l)^2+B, \text{ for } b \leqslant x \leqslant \infty$$

$$\text{with } b = B/2kl,$$

since the splitting of the energy levels is not critically dependent on the shape of the potential function; $2l$ is the distance between the two minima, $k = (2\pi\nu_0)^2 \cdot m$ is a force constant. The unsymmetry parameter B, which is a measure of the dipole–dipole interaction, equals $2\mu\mathscr{F}$. Here, μ is the component of the dipole moment in the direction of the local field:

$$\mathscr{F} = \mathscr{E}+f\mathscr{P}.$$

* In the local field approximation, the direction of P_d is not important.

The other symbols used have their usual significance.

As it has been shown elsewhere,[6] in the case of potential barriers that are not too low, approximative eigenvalues of the resulting wave equation may be found by using the variational method. Acceptable wave functions are obtained by linear combination of harmonic oscillator eigenfunctions. The resulting energy-level expressions[6] are rather tedious to handle, but since we are primarily interested in the energy levels near the bottom of the barrier, further simplifications may be introduced. By putting the overlap integrals equal to zero and taking

$$H_{11}^{1} = \int \psi_1 H \psi_1 \, dx = E_1^0, \quad H_{22}^{2} = \int \psi_2 H \psi_2 \, dx = E_2^0,$$

we get:

$$E_{n\pm} = \bar{E}_n \pm \lambda_n \beta_n, \tag{1}$$

where $\bar{E}_n = \frac{1}{2}(E_1^0 + E_2^0)$ is the mean of the unperturbed levels and λ_n equals:

$$\lambda_n = \left[1 + \frac{\mu^2 \mathscr{F}^2}{\beta_n^2}\right]^{1/2}$$

with

$$\beta_n = \int \psi_{1,n} H \psi_{2,n} \, dx.$$

The coefficients of the harmonic oscillator wave functions ψ_1 and ψ_2 are:

$$C_1 = \left[\frac{(4\beta^2 + B^2)^{1/2} + B}{2(4\beta^2 + B^2)^{1/2}}\right]^{1/2}$$

and

$$C_2 = \left[\frac{(4\beta^2 + B^2)^{1/2} - B}{2(4\beta^2 + B^2)^{1/2}}\right]^{1/2}$$

First the spectral consequence of this model will be discussed. Equation (1) shows that, owing to the tunnel effect, the protonic vibrational levels are split into doublets. As the result of the transitions $n_{0+} \leftrightarrow n_{1-}$ and $n_{0-} \leftrightarrow n_{1+}$, two OH stretching bands appear instead of one in the 1800–2800 cm^{-1} region. According to the selection rules, the splitting of the ν_{OH} bands should amount to $\Delta E_0 + \Delta E_1$ in the i.r. and to $\Delta E_1 - \Delta E_0$ in the Raman spectrum. In agreement with this prediction, a larger band splitting was always found in the i.r. than in the Raman spectra.

As to the p.m.r. spectra, the proposed model predicts a protonic motion with a correlation frequency:

$$\nu_r = \frac{\sum i \Delta \nu_i, ef \exp(-\bar{E}_i/kT)}{\sum i \exp(-\bar{E}_i/kT)}.$$

This motion becomes temperature-independent at low temperatures as the result of tunnelling of the proton on the lowest vibrational level. The experiments have confirmed this fact.

Since it has been shown that the proposed model is consistent with the spectral data, we proceed to calculate the polarization properties. The partition function in the presence of the local field $\mathscr{F}$ is:

$$Z = 2\cosh(\lambda_0 \beta_0/kT) + \\ + \sum_{n=1}^{\infty} \exp(-\bar{E}_n/kT) \cdot 2\cosh(\lambda_n \beta_n/kT). \tag{2}$$

From this we get the free energy of the system as:

$$A_E = -NkT \ln Z.$$

On the other hand we may, following DEVONSHIRE, expand $A_E = A - A_0$ in powers of P:

$$A_E = \tfrac{1}{2}aP^2 + \tfrac{1}{4}bP^4 + \ldots, \tag{3}$$

the coefficients a, b, ... being functions of the temperature.

The coefficient a, which determines the ferroelectric transition point, can be obtained from our model. Observing the condition, that in the vicinity of $P = 0$ the local field $\mathscr{F}$ will be vanishingly small, and using the expressions

$$\mathscr{E} = \mathscr{F} - f(\mathscr{P}_d + \gamma \mathscr{F}), \quad \mathscr{P}_d = -(\partial A_E/\partial \mathscr{F})_T$$

$$a = (\mathscr{E}/\mathscr{P})_{\mathscr{P}=0}$$

we get a as:

$$a = \frac{\dfrac{1-\gamma f}{\mu^2 f N} - \dfrac{\left[\sinh(\beta_0/kT) \cdot \dfrac{1}{\beta_0} + \sum_{n-1}^{\infty} \exp(-\bar{E}_n/kT) \cdot \sinh(\beta_n/kT)\dfrac{1}{\beta_n}\right]}{\left[\cosh(\beta_0/kT) + \sum_{n=1}^{\infty}\exp(-\bar{E}_n/kT) \cdot \cosh(\beta_n/kT)\right]}}{\dfrac{\gamma}{\mu^2 f N} + \dfrac{\dfrac{1}{f}\left[\sinh(\beta_0/kT) \cdot \dfrac{1}{\beta_0} + \sum_{n=1}^{\infty} \exp(-\bar{E}_n/kT) \cdot \sinh(\beta_n/kT) \cdot \dfrac{1}{\beta_n}\right]}{\left[\cosh(\beta_0/kT) + \sum_{n=1}^{\infty} \exp(-\bar{E}_n/kT) \cdot \cosh(\beta_n/kT)\right]}} \tag{4}$$

It may be pointed out that the change of the bond moment with the local field does not enter into the expression for a, as $d\mu/d\mathscr{F}$ occurs only multiplied with $\mathscr{F}^2$ and remains finite as $\mathscr{F} \to 0$. In deriving equation (4) we have neglected terms containing $\partial\beta/\partial\mathscr{F}$, as β is a slowly changing function of $\mathscr{F}$.

In most cases we may disregard in equation (3) terms with $n \geqslant 1$. Putting $\gamma = 0$, the expression for a simplifies into:

$$a = f[\beta_0 f\mu^2 N \cdot \coth(\beta_0/kT) - 1]$$

In the classical limit ($\beta \to 0$) a reduces to a linear function of the temperature:

$$a = (-f\mu^2 N + kT)/N\mu^2$$

which was obtained already by DEVONSHIRE and MASON.

In this case the expression $(\partial a/\partial T)^{-1}$ equals the constant in the Curie–Weiss law: $C = N\mu^2/k$. In the quantum case, however, $(\partial a/\partial T)^{-1}$ becomes temperature- and mass-dependent and equals to:

$$(\partial a/\partial T)^{-1} = C \cdot \sinh^2(T_0/T) \cdot (T/T_0)^2$$

with $T_0 = \beta/k$.

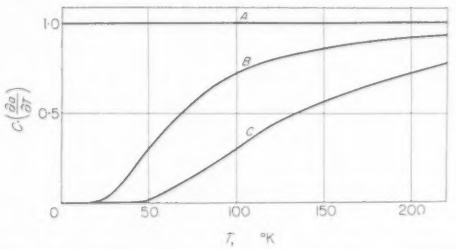


FIG. 1. $C(\partial a/\partial T)$ as a function of temperature for: (A) $T_0 = 0$°K, (B) $T_0 = 100$°K and (C) $T_0 = 200$°K

The temperature T_0 divides the quantum from the classical region. If a crystal undergoes a transition to the ferroelectric state at a temperature above T_0, the quantum effects are small.

3. RESULTS

From the condition $a = 0$ we get for the Curie temperature T_C the following equation:

$$\frac{\beta_0}{f\mu^2 N}(1-\gamma f) = \frac{\sinh(\beta_0/kT_C) + \sum_{n=1}^{\infty} \exp(-\bar{E}_n/kT_C) \cdot \sinh(\beta_n/kT_C) \cdot (\beta_0/\beta_n)}{\cosh(\beta_0/kT_C) + \sum_{n=1}^{\infty} \exp(-\bar{E}_n/kT_C) \cdot \cosh(\beta_n/kT_C)}. \tag{5}$$

VOL. 13 1960

which is best solved graphically. Choosing reasonable values for $(1-\gamma f)/f\mu^2 N$, we find that in the case of KH_2PO_4 the value of $hc\beta$ must be smaller than 200 cm^{-1} in order to get real roots for the Curie temperature.

The same condition remains valid also in the case of anharmonic single-minimum potential fields, for which we obtain similar Curie-point expressions. As in these cases we have no vibrational splitting, β_0 ought to be equal to $\frac{1}{2}\nu_{OH}$. Anharmonic single-minimum models therefore predict the occurrence of the ν_{OH} band in the region of ~400 cm^{-1}. However, this is at variance with the measured i.r. data: ν_{OH}~1800–2800 cm^{-1}. On the other hand the condition $\beta_0 = \Delta E_0/2 \ll \nu_{OH}$ fits well in the double-minimum model and is consistent with the measured doubling of the OH stretching bands and the n.m.r. data. So the possibility of a single-minimum potential field may be definitely rejected.

In the quasi-classical region ($\beta < kT_C$) we get, neglecting γ, T_C by series expansion explicitly as:

$$T_C = (K/k)[1-(\tfrac{1}{3})(\beta/K)^2] \qquad (6)$$

with $K = fN\mu^2$. The first term K/k, which does not depend on the mass of the vibrator, is identical with the classical expression of MASON and DEVONSHIRE. Since the splitting of the zeroth vibrational level diminishes on deuteration, owing to the greater difficulty of barrier penetration by the deuteron, we get the correct dependence of the Curie point on the mass of the hydrogen isotope:

$$T_{C_D} \geqslant T_{C_H}.$$

Of course the result $T_{C_D} > T_{C_H}$ remains valid in the quantum region too. In the limit $\beta \gg kT_C$, we obtain T_C approximately as:

$$T_C = -2\beta/k \ln \tfrac{1}{2}(1-\beta/f\mu^2 N) \qquad (7)$$

On rearrangement, this becomes $z = -\alpha/\ln\frac{1}{2}(1-\alpha)$ with $z = kT_C/2fN\mu^2$ and $\alpha = \beta/f\mu^2 N$. Fig. 2 shows the curve $z = z(\alpha)$ in the range $0{\cdot}9 < \alpha < 1$. The limiting value of z is obtained from equation (5) as: $\lim_{\alpha \to 0} z = 0{\cdot}5$.

The proposed model predicts a Curie-point ratio T_{C_D}/T_{C_H} between 1·66 and 1 for the region $0 \leqslant \alpha \leqslant 0{\cdot}9$, which is most likely to occur. For the region $0{\cdot}90 \leqslant \alpha \leqslant 0{\cdot}99$ our model predicts a ratio

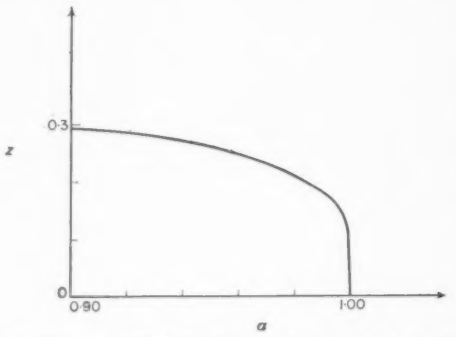


FIG. 2. $z = kT_C/2fN\mu^2$ as a function of α in the "quantum" region.

T_{C_D}/T_{C_H} between 1·66 and 2. Only for the exceptional case $0{\cdot}99 \leqslant \alpha \leqslant 1$ would this ratio exceed the value 2. As can be seen from Table 1, the experimental data seem to confirm this prediction.

Table 1

	T_{C_H}(°K)	T_{C_D}(°K)	T_{C_D}/T_{C_H}
KH_2PO_4	122	213	1·74
KH_2AsO_4	96	162	1·69
RbH_2PO_4	146	218	1·49
RbH_2AsO_4	110	178	1·62
CsH_2PO_4	159	?	?
CsH_2AsO_4	143	212	1·48
$NH_4H_2PO_4$	147	229	1·56
$NH_4H_2AsO_4$	216	299	1·38
$Ag_2H_3IO_6$	225	279	1·24
$(NH_4)_2H_3IO_6$	254	266	1·05

In agreement with equation (5) we see, too, that higher Curie temperatures, i.e. greater dipole–dipole interactions, result in smaller isotopic effects.

In agreement with equation (5), equation (7) shows that the Curie temperature approaches zero for $\alpha \to 1$. This means, that the ferroelectric transition occurs only if the dipole–dipole interactions, expressed as frequencies, are greater than the tunnelling frequency on the zeroth vibrational level. As the tunnelling frequency diminishes on deuteration, cases may exist where the ferroelectric transition takes place only for the deuterated, but not for the undeuterated, compound.

Physically, we may interpret the higher Curie point T_{C_D} as the result of the greater dipole–dipole

VOL. 13 1960

interactions due to the greater effective dipole moment in the deuterated compounds. The existence of a greater effective dipole moment for deuterated than for hydrogen compounds is implicitly contained in our energy-level expression.

From

$$-\boldsymbol{M} = (\partial E/\partial \boldsymbol{F}_x, \partial E/\partial F_y, \partial E/\partial F_z)$$

we get:*

$$|M/M_0| = (B/2\beta)/[1+(B/2\beta)^2]^{1/2} \qquad (8)$$

The same result may be more intuitively derived from $M = M_0[C_1^2 - C_2^2]$, if we remember that the effective dipole moment is proportional to the difference between the probabilities of finding the proton in the well 1 or well 2.

It may be seen from Fig. 3 that the effective dipole moment M is a function of the ratio between the interaction of the classical dipole moment with

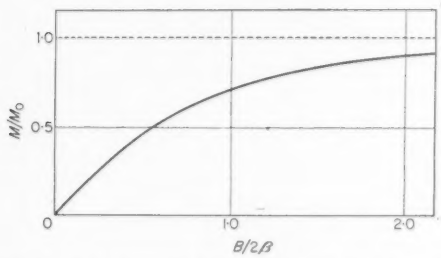


FIG. 3. The effective dipole moment as a function of $B/2\beta$.

the local field and the splitting of the zeroth vibrational level. In the case of long hydrogen bonds, we have a large potential barrier and therefore $\beta \to 0$, $\beta \ll B$. Here M approaches its classical value. The transition to the ferroelectric state is expected to occur at relatively high temperatures. On the other hand we have in the case of very short hydrogen bonds a vanishingly small barrier $\beta \gg B$, so that M approaches zero. Then, of course, no ferroelectric transition occurs.

In the intermediate case M becomes mass-dependent: $M_0 \geqslant M_D \geqslant M_H$.

Something must be said about the role of the Ubbelohde effect with short hydrogen bonds. The experimental Curie-point data for KH_2PO_4 and KD_2PO_4 might be fitted by equation (7) by taking $\alpha_H = 0{\cdot}9$ and $\beta_H = 107\ cm^{-1}$. But in this case the splitting $2\beta_H$ would become incompatible with the assignment of the i.r. spectra of these compounds.

The correct explanation would take into account the fact that each well of a short double-minimum hydrogen bond is separately strongly anharmonic, and that the hydrogen bonds expand on deuteration[12] thus raising the classical dipole moment, if the condition $\partial\mu/\partial r > 0$ is satisfied. This effect, which is itself a consequence of the quantization of the protonic motion,[13] increases the quantum effects, thus further raising the Curie point of the deuterated compound. By including this we obtained good agreement between the theoretically calculated and the experimental Curie points for KH_2PO_4 and KD_2PO_4.

In the case of long hydrogen bonds the barriers are large and the splitting of the zeroth level becomes very small. In addition, both wells become increasingly harmonic, so there is no expansion of the hydrogen bonds on deuteration. As a result of these two facts, the isotopic effects are small, as found by experiment.

The temperature-dependence of the spontaneous polarization is given by:

$$P_d = -(\partial A_E/\partial \mathscr{F})_T = (N\mu/\lambda)(\lambda^2-1)^{1/2}[1+(\mathscr{F}/\mu)(\partial\mu/\partial\mathscr{F})]\tanh(\lambda_0\beta_0/kT). \qquad (9)$$

Neglecting the term $\partial\mu/\partial\mathscr{F}$, introducing $y = P_d/N\mu$, $x = T/T_C$, and using for T_C equation (6), we can rewrite this equation in a more convenient form:

$$(\alpha^2+y^2)^{1/2} = \tanh\,[(\alpha^2+y^2)^{1/2}/x(1-\tfrac{1}{3}\alpha^2] \qquad (10)$$

For the classical case, where $\alpha = 0$, equation (10) goes over into that of DEVONSHIRE and MASON: $y = \tanh y/x$.

Since we have used the quasi-classical equation (6), equation (10) also holds strictly only for the region $\alpha \ll 1$. Nevertheless, some conclusions can be drawn from this equation which are valid quite generally.

* M_0 is the classical dipole moment. For a more detailed discussion see: S. BRATOŽ, *J. Chim. Phys.*, **56**, 260 (1959).

We find that in the region $x>1$, i.e. above the transition temperature, the only value of y to satisfy equation (10) is zero. Only below that

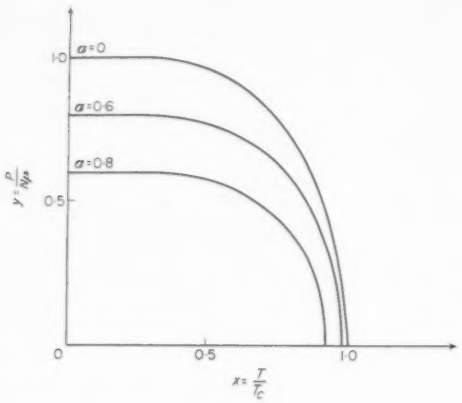


FIG. 4. Normalized spontaneous polarization as a function of the reduced temperature according to equation (10).

temperature does there exist a non-zero solution. Further we see that, in agreement with equation (4), a real solution for y exists only for $0\leqslant\alpha\leqslant1$. This means that a ferroelectric transition occurs only when the electrostatic interactions are not too small in comparison with the zero-level splitting.

As the result of the quantization of the protonic motion, the polarization curves for different α values are distinctly separated. Since α diminishes on deuteration, the KD_2PO_4 curve now lies above the curve for KH_2PO_4, as revealed by ZWICHER and SCHERRER[14]. For the same reason the value of the spontaneous polarization at absolute zero is greater for deuterium than for hydrogen crystals. It is interesting to note that as long as α differs from zero, y can never reach its saturation value.

In deriving our model, we have, following DEVONSHIRE and MASON, considered only long-range forces. The true model, which should take into account also short-range interactions, would give, in agreement with experiment, rather sharper transitions.

From the proposed model we may further derive the entropy S of the transition, corresponding to the disorder of the hydrogen positions. Neglecting terms with $n\geqslant1$, we obtain:

$$S=-(\partial A_E/\partial T)_{\mathscr{F}}$$
$$=Nk\{\ln[2\cosh(\lambda x)]-(\lambda x)\tanh(\lambda x)\} \qquad (11)$$

with $x=\beta/kT$.

In the same approximation the extra specific heat of the transition is obtained from $C=T\partial S/\partial T$. Here too, the isotopic effects have the correct sign:

$$\Delta C_D\geqslant\Delta C_H.$$

4. CONCLUSIONS

An attempt to correlate the isotopic effects in the ferroelectric properties of hydrogen-bonded crystals with their i.r. and n.m.r. spectra has led to the conclusion that the protons in the hydrogen bonds are tunneling in double-minimum potential fields. The ferroelectric transition is assumed to be the result of a deformation of the protonic distribution due to the electrostatic interactions. As a first approximation, DEVONSHIRE's and MASON's theory of ferroelectricity in KH_2PO_4 was extended by treating the problem quantum-mechanically instead of classically. It was found that the ferroelectric transition occurs only if the dipole–dipole interactions expressed as frequencies are greater than the tunnelling frequency on the lowest vibrational level. The proposed model is able to explain on the same basis:

(1) The existence of two ν_{OH} bands instead of one in the i.r. and Raman spectra, the splitting being smaller in the Raman than in the i.r.

(2) The fact that both ν_{OH} bands change only a little on going through the transition point, in spite of great apparent changes in OH bond lengths.

(3) The existence of rapid protonic motions in the p.m.r. spectra, which become temperature-independent at low temperatures.

(4) The dependence of the ferroelectric Curie point on the mass of the hydrogen isotope.

(5) The sharper increase of the spontaneous polarization with falling temperature for the deuterated compound.

(6) The greater value of the spontaneous polarization at absolute zero for deuterated than for hydrogen compounds.

The necessity of taking into account the hydrogen-bond expansion on deuteration in the case of the short hydrogen bond has been emphasized. Further work, including the study of

short-range and electromechanical forces, seems to be desirable to explain the ferroelectric phenomena in their whole complexity.

Acknowledgement—The author is very indebted to Professor D. HADŽI, Professor A. PETERLIN and Dr. S. BRATOŽ for their continued interest and many helpful discussions, as well as to Professor I. KUŠČER, who read the manuscript critically.

REFERENCES

1. SLATER J. C., *J. Chem. Phys.* **9**, 16 (1941).
2. MASON W. P., *Phys. Rev.* **72**, 854 (1947).
3. DEVONSHIRE A. F., *Advanc. in Phys.* **3**, 85 (1954).
4. GRINDLEY J. and TER HAAR D., *Proc. Roy. Soc.* A**250**, 267 (1959).
5. PIRENNE J., *Physica s'Grav.* **15**, 1019 (1949).
6. BLINC R. and HADŽI D., *Molecular Phys.* **1**, 391 (1958).
7. REID C., *J. Chem. Phys.* **30**, 182 (1959).
8. BACON G. E. and PEASE R. S., *Proc. Roy. Soc.* A**230**, 359 (1955).
9. BADGER R. M., *J. Chem. Phys.* **2**, 128 (1934).
10. NEWMAN R., *J. Chem. Phys.* **18**, 669 (1950).
11. BETHE H., *Proc. Roy. Soc.* A**150**, 552 (1935).
12. UBBELOHDE A. R. and GALLAGHER K. J., *Acta Cryst., Camb.* **8**, 71 (1955).
13. NORDMANN C. E. and LIPSCOMB W. N., *J. Chem. Phys.* **19**, 1422 (1951).
14. ZWICHER B. and SCHERRER P., *Helv. Phys. Acta* **17**, 346 (1944).

VOL.
13
1960

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 13. pp. 212–220. Printed in Great Britain.

MAGNETIC SUSCEPTIBILITY, HEAT CAPACITY AND THIRD-LAW ENTROPY OF $MgNi_2$*†

J. S. WOLLAM‡ and W. E. WALLACE

Department of Chemistry, University of Pittsburgh, Pittsburgh 13, Pennsylvania

(*Received* 11 *September* 1959)

Abstract—A study of the heat capacity and magnetic susceptibility of $MgNi_2$ was undertaken to evaluate the report by Voss that this compound is ferromagnetic at temperatures below 235°C. The results of the investigation reveal no evidence of ferromagnetism in this Laves phase. $MgNi_2$ is found to be a moderately strongly paramagnetic material with a susceptibility which is practically constant between 25 and 450°C. Thermomagnetic analysis shows a very slight but abrupt drop in the susceptibility at 194°C, and heat-capacity measurements reveal a very small thermal anomaly at the same temperature. The effects are thought to be associated with some type of irregularity in the lattice, such as the defect structure which occurs when a nickel atom replaces a magnesium atom in the $MgNi_2$ lattice. Third-Law entropies of $MgNi_2$ at temperatures up to 555°K are also presented.

1. INTRODUCTION

THE phase diagram of the magnesium–nickel system is based to a considerable extent on results published by VOSS[1] in 1908. At that time there was reported the existence of a discrete phase which contained 82·84 per cent nickel, corresponding to the formula $MgNi_2$, together with the further claim that the phase was ferromagnetic, with its Curie point at 235°C. The existence of the compound was subsequently confirmed by LAVES and WITTE[2], who used X-ray diffraction techniques to work out its structure.

The reputed ferromagnetism of $MgNi_2$, if confirmed, would be of considerable interest because of its implications with regard to the electronic status of nickel in this compound. It is a well-known fact that as nickel is alloyed with any of a variety of other metals, its saturation magnetization is progressively reduced as the concentration of the second metal increases. One example of this is the nickel–copper system, in which ferromagnetism is not found in the solid solutions which contain copper in excess of 60 atomic per cent. With zinc as the solute the rate of decrease of magnetization with concentration indicates that quenching of the ferromagnetism would occur at a zinc content of about 30 atomic per cent if the range of primary solubility extended that far. These effects presumably result from the absorption of the extra electrons, supplied by the solute atoms, into the partially vacant nickel *d*-orbitals. Consequently, the magnetic moment decreases with increasing solute concentration until at some point the *d*-orbitals are filled, or nearly so, and the solid is rendered non-ferromagnetic.

On this basis, it seems strange that $MgNi_2$ was found to be ferromagnetic, since nickel in this compound is in chemical union with $33\frac{1}{3}$ per cent divalent metal. Structurally, however, the compound is very different from metallic nickel, so that the crystal field in the compound undoubtedly differs markedly from that occurring in the nickel-based primary solid solutions which contain copper or zinc as solute. It would therefore seem possible that the behavior of nickel in the compound is sufficiently altered to permit some of the *d*-orbitals to remain empty despite the high concentration of

* From a thesis submitted by J. S. WOLLAM to the University of Pittsburgh in partial fulfillment of the requirements for the Ph.D. degree, January, 1958.

† This work was assisted by the U.S. Atomic Energy Commission.

‡ Now at the Knolls Atomic Power Laboratory of the General Electric Company, Schenectady, New York.

VOL. 13 1960

quasi-free electrons. If the possibility of such an occurrence be admitted, the ferromagnetism reported by Voss does not seem so implausible.

Aside from its reputed ferromagnetism, there is another respect in which $MgNi_2$ seems to be aberrant. A large number of intermetallic compounds whose overall composition can be represented by the formula AB_2 crystallize in one or another of three closely-related structures. These are the $MgCu_2$, $MgNi_2$ and $MgZn_2$-type structures, which as a group are known as the Laves phases. Such compounds are thought to occur primarily because space may be conveniently and densely filled by a lattice of one of these types when the ratio of the atomic diameters of A and B is about $\sqrt{3}/2$, or 1·225. However, LAVES and WITTE[3] have shown that electronic factors are also involved, in that the electron:atom ratio appears to determine which of the three structures a given compound will adopt. Valence-electron concentrations of 1·3–1·8, 1·8–1·9 and 1·9–2·1 seem to favor the $MgCu_2$, $MgNi_2$ and the $MgZn_2$ structures, respectively. In the compound $MgNi_2$ itself, the electron:atom ratio would be 0·67 if the nickel atoms were zero valent and would be even less if the nickel-ion cores absorbed valence electrons, as they have been shown to do in some instances. On the other hand, ferromagnetism in the compound would imply incomplete filling of the *d*-orbitals in the nickel. If the intensity of magnetization were great enough, it would perhaps indicate an unpairing and transfer into the valence band of a sufficient quantity of *d*-electrons to bring the valence-electron concentration into the expected range.

It was for the purpose of establishing whether $MgNi_2$ is actually ferromagnetic and, further, to provide information about the electronic character of nickel and the electron concentration in this compound, that the magnetic studies were carried out. The thermal measurements were made partly with the expectation of establishing the magnetic energy of the system and partly to enable the evaluation of the Third-Law entropy of the compound to be made. As will be seen later, the claim by Voss that $MgNi_2$ is ferromagnetic was not substantiated, but the study did reveal the existence of a small magnetic effect in $MgNi_2$ at 194°C which has not heretofore been reported in the literature.

2. EXPERIMENTAL DETAILS

(a) *Sample preparation*

The samples used in the work were prepared by fusing the elements together in a high-purity MgO crucible under an atmosphere of purified argon. The magnesium had been supplied without charge by the Dow Chemical Company and was received in the form of a polycrystalline slab. It had been refined by sublimation until its purity exceeded 99·95 per cent, with any contaminants present as spectroscopic traces only. The nickel was donated by the International Nickel Company,* who stated that its purity exceeded 99·9 per cent. The principal impurities were oxygen, carbon and cobalt, which were present to the extent of 0·065, 0·014 and 0·0018 per cent, respectively.

To form $MgNi_2$ it is necessary to heat the mixed metals to about 1200°C. Because of the volatility of magnesium at this temperature, which exceeds its normal boiling point, attempts to prepare samples in which the elements were present in precisely the correct stoichiometric ratio were never successful. The samples used for the magnetic-susceptibility measurements were slightly rich in magnesium (81·87 per cent nickel instead of the theoretical 82·84 per cent), whereas the sample used for determining heat capacities contained an excess of 0·58 weight per cent nickel.† Excess magnesium was avoided in the heat-capacity sample, since this would have resulted in the formation of some Mg_2Ni, whose heat capacity is unknown. In this part of the work an excess of nickel was preferable, since its heat capacity is well known and an appropriate correction could be made. On the other hand, the presence of excess nickel, or any other ferromagnetic contaminant, was highly undesirable in the samples used in the susceptibility determinations.

Spectroscopic examination of the $MgNi_2$ showed iron, silicon and calcium as the principal metallic impurities, probably introduced by the crucible. They were each present to the extent of ~0·01 per cent. Smaller amounts of aluminum, chromium and molybdenum were also present, but the total amount of metallic impurities was less than 0·05

* The assistance of Dr. W. A. WESLEY in arranging this donation is gratefully acknowledged.

† The nickel content of the samples was determined by the usual electrolytic method.

per cent. The position and intensities of the lines in X-ray diffraction patterns for the substance were in excellent agreement with those reported by LAVES and WITTE[2] for $MgNi_2$.

In the early stages of the work attempts were made to improve the homogeneity of the sample by heat treating at 800–1000°C. This was entirely unsuccessful because of the volatility of magnesium. For this reason all measurements were made on samples which had been allowed to solidify in the crucible after fusion by induction heating. When full advantage was taken of the magnetic stirring inherent in this mode of heating, it was found that the resulting ingot was uniform in composition to within 0·2 per cent, so that no further treatment was needed.

(b) *Thermal measurements*

The equipment used and the procedure employed to determine the heat capacities have been described in earlier publications from this Laboratory.[4, 5] Two calorimeters were employed—one for temperatures below, the other for temperatures above, 300°K. The sample weights for the upper and lower temperature ranges were 338·676 and 338·004 g, respectively.

(c) *Magnetic measurements*

Susceptibility measurements were made by the method of Faraday, the force being measured by means of an Ainsworth analytical balance. Two types of determination were made. One consisted in scanning the temperature range (thermomagnetic analysis) to ascertain whether the relative susceptibility of the sample might show a drop which could be ascribed to the Curie point for $MgNi_2$. The other type was of such a nature as to permit evaluation of the true susceptibility of the material.

A six-inch Varian electromagnet was used to establish the magnetic field, with the magnetizing current supplied by a motor-generator set. A control circuit similar to that described by HEDGECOCK and HUNT[6] had been applied to the generator field windings to obtain a current constant to within about 0·1 per cent. Constant-gradient pole caps were used to provide a Sucksmith gap. In the thermomagnetic analysis, which required the use of a furnace, the gap width was 1 in. and the maximum field strength was about 14 kOe. In the other type of measurement the gap was closed down to $\frac{1}{2}$ in. and a maximum field of 21 kOe was attained.

In most instances the Pyrex tubes, containing the sample in the form of a fine powder, were sealed off under helium. In a few cases the samples were sealed off under hydrogen and then heat treated. This was done because Voss had used hydrogen as an inert atmosphere in preparing his alloys, so that there was a possibility that ferromagnetism had been induced by hydrogen uptake by the samples.

The true susceptibility of $MgNi_2$ was determined by measuring the apparent susceptibilities as a function of field strength and extrapolating to infinite field by the method of HONDA[7] to eliminate the effect of ferromagnetic impurities. The field-strength measurements needed to make such a plot were obtained with the aid of a model D79 Dyna-Empire gaussmeter. The field gradient was determined using both $NiCl_2$ solution and Mohr's salt, following procedures suggested by SELWOOD[8].

3. RESULTS

(a) *Magnetic measurements*

The experimental data provide no support for the claim by Voss that $MgNi_2$ is ferromagnetic. Results of the thermomagnetic analysis, shown in Fig. 1, give no indication of a Curie point at 235°C. The fall in susceptibility at about 350°C is almost certainly due to the presence of a small amount of elemental nickel, presumably there because of slightly nickel-rich regions in the sample.* The decrease is much too small to be ascribed to $MgNi_2$ as a whole. The amount of elemental nickel indicated from the results in Fig. 1 is 0·07 per cent. The quantity of ferromagnetic impurity can also be estimated from the slope of the Honda plot. From this, one also obtains 0·07 per cent as the amount of uncombined nickel, supporting the contention that the Curie point observed at 350°C is to be ascribed to nickel, rather than to $MgNi_2$.

The mass susceptibility of $MgNi_2$ at 25°C is found from the Honda plot to be $2{\cdot}80\times10^{-6}$ e.m.u./g. Attempts to employ a similar procedure

* Diffraction patterns of the nickel-rich sample used for the heat-capacity determinations showed faint lines characteristic of elemental nickel, indicating that $MgNi_2$ will not hold excess nickel in solid solution.

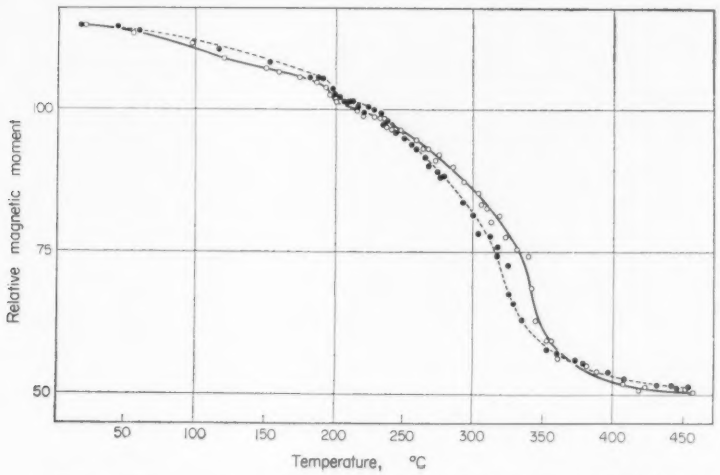


FIG. 1. Thermomagnetic analysis of $MgNi_2$ ○–○ untreated ●–● hydrogenated.

to obtain values between 25 and 360°C were unsuccessful, since the furnace used to reach elevated temperatures necessitated widening the magnet air gap to 1 in. Under these circumstances the maximum attainable field was insufficient to permit a reliable extrapolation to infinite field. However, above the nickel Curie point the susceptibility could again be determined. It was found to be $2{\cdot}69\times10^{-6}$ e.m.u./g at 450°C after correction for the known amounts of elemental nickel and magnesium in the sample.

A small break in the thermomagnetic-analysis curves at about 194°C is apparent in Fig. 1. This represents an abrupt drop in the mass susceptibility of about $0{\cdot}15\times10^{-6}$ e.m.u./g, the origin of which is not clear. The effect is much too small to attribute to the destruction of ferromagnetism in $MgNi_2$. It is nevertheless a real effect, since, as will be pointed out presently, a slight anomaly in the heat capacity occurs at this same temperature.*

* *Note added in proof:* It has now been brought to the author's attention that WITTE and co-workers at the University of Darmstadt have also observed an anomaly in the thermal expansivity at about 194°C.

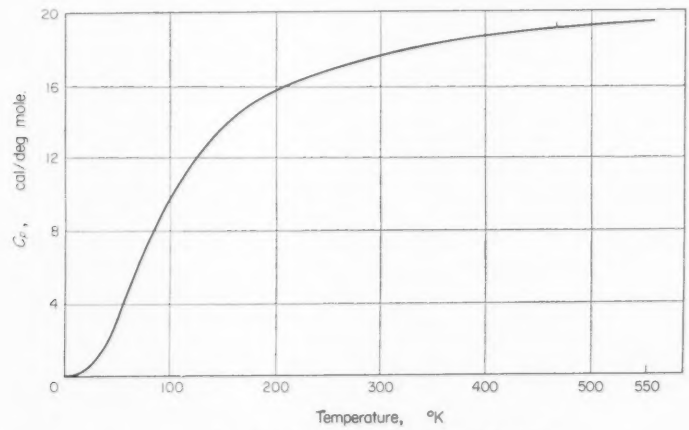


FIG. 2. Molar heat capacity of $MgNi_2$.

(b) *Heat capacities*

Results of the heat-capacity determinations are shown in Table 1. Judging from the scatter of the data from a smooth curve, they are precise to within $\pm 0{\cdot}1$ per cent except for temperatures below 50 and above 500°K. At the highest temperatures the scatter averages about 0·25 per cent whereas at the lowest temperatures it rises to about 1 per cent.

The variation of C_p with temperature shown in Fig. 2 is as expected, except in the range between 460 and 470°K. An expanded plot of the data for this region is given in Fig. 3. There is no indication of a λ point at 508°K, the Curie temperature indicated by Voss for $MgNi_2$. There is, however, a small but clearly recognizable thermal anomaly at about 465°K. This is the temperature at which there was a break in the thermomagnetic curve, indicating that the excess heat capacity and the magnetic effect have a common origin.

The only previous study of the heat capacity of $MgNi_2$ was that made by SCHÜBEL[9] many years ago. He claimed a precision of about 2 per cent for his results. Comparison of the two sets of data indicates agreement to within about 2 per cent over the temperature range common to both the earlier drop calorimetry and the present adiabatic work.

The change in the heat capacity associated with

Table 1. *Heat capacities of* $MgNi_2$

Temperature (°K)	C_p (cal. deg.$^{-1}$ mole^{-1})	Temperature (°K)	C_p (cal. deg.$^{-1}$ mole^{-1})
Series I*		Series I (continued)	
10·55	0·0403	136·90	12·836
11·77	0·0487	142·39	13·176
13·10	0·0585	147·85	13·494
14·90	0·1061	153·25	13·794
17·32	0·1544	158·65	14·068
19·46	0·2186	164·09	14·325
21·37	0·2918	169·64	14·582
23·24	0·3794	175·31	14·805
25·11	0·4857	181·18	15·034
27·29	0·6324	187·21	15·244
27·63	0·6536	193·22	15·453
31·19	0·9481	199·29	15·638
34·21	1·218	209·79	15·918
36·68	1·488	216·07	16·092
38·72	1·713	217·80	16·128
41·02	1·983	222·33	16·237
43·93	2·358	224·02	16·281
48·39	2·952	230·27	16·433
53·54	3·684	236·53	16·577
58·29	4·364	242·82	16·691
63·44	5·125	249·09	16·818
69·46	5·968	255·35	16·930
74·66	6·698	261·61	17·028
80·43	7·470	267·90	17·134
86·16	8·216	274·21	17·211
92·40	8·957	280·54	17·318
98·30	9·592	286·90	17·407
103·93	10·163	293·29	17·485
109·40	10·692	287·09	17·390
114·73	11·172	293·47	17·476
120·03	11·602	299·83	17·576
125·62	12·052	306·18	17·641
131·34	12·460	312·50	17·709
		318·81	17·789

Table 1. *Heat capacities of* $MgNi_2$ (Contd.)

Temperature (°K)	C_p (cal. deg.$^{-1}$ mole^{-1})	Temperature (°K)	C_p (cal. deg.$^{-1}$ mole^{-1})
Series II*		Series II (continued)	
306·92	17·71	473·07	19·04
308·74	17·66	478·76	19·06
312·83	17·76	484·44	19·10
314·67	17·77	490·11	19·14
320·56	17·85	495·78	19·17
326·45	17·91	501·44	19·20
332·33	17·98	507·12	19·25
338·18	18·02	512·79	19·34
344·02	18·11	518·45	19·28
349·84	18·15	524·11	19·23
360·54	18·24	529·77	19·40
366·34	18·30	535·44	19·33
372·12	18·35	541·11	19·34
374·73	18·36	546·75	19·43
380·51	18·42	552·36	19·42
386·27	18·45	558·03	19·35
392·03	18·53		
397·78	18·53	Series III	
403·51	18·64	310·68	17·72
409·25	18·68	319·55	17·81
414·97	18·69	328·37	17·91
420·68	18·71	461·76	18·98
426·38	18·76	465·92	20·11
432·08	18·81	469·63	18·99
437·76	18·85	511·63	19·28
443·46	18·88	517·22	19·34
449·15	18·89	522·83	19·34
454·83	18·93	528·42	19·34
456·10	18·94	534·01	19·38
461·78	18·96	539·59	19·43
467·43	19·36	545·17	19·39

* The Series I measurements were made in the low-temperature calorimeter. The others were made in the high temperature calorimeter. The measurements in Series III repeat a portion of those in Series II, partly for the improvement of precision but mainly for confirmation of the existence of the thermal anomaly at 194°C.

VOL. 13 1960

formation of the compound from the elements (ΔC_p), or the deviation from the Kopp–Neumann additivity rule for heat capacities, is given in Table 2. It is seen that at all temperatures above 50°K, ΔC_p is negative. The deviations from additivity show in general a uniform increase with temperature, ranging from 3·3 per cent at 50°K to 14·5 per cent at 550°K. This behavior implies a considerable change in the type or strength of bonding when the elements unite to form the compound. There is, on the average, less freedom of atomic motion in the compound than in the elements, a conclusion which is also indicated by the exothermal formation of the material, and by its brittle hardness.

(c) *Entropies*

The heat-capacity data were numerically integrated in the usual way to obtain Third-Law entropies of $MgNi_2$. Extrapolation below 10°K was made by assuming a simple T^3 law. The corresponding entropy for the region of extrapolation was 0·012 e.u. per mole of $MgNi_2$. The Third-Law entropies and entropies of formation for several temperatures are given in Table 3. For calculation of the latter

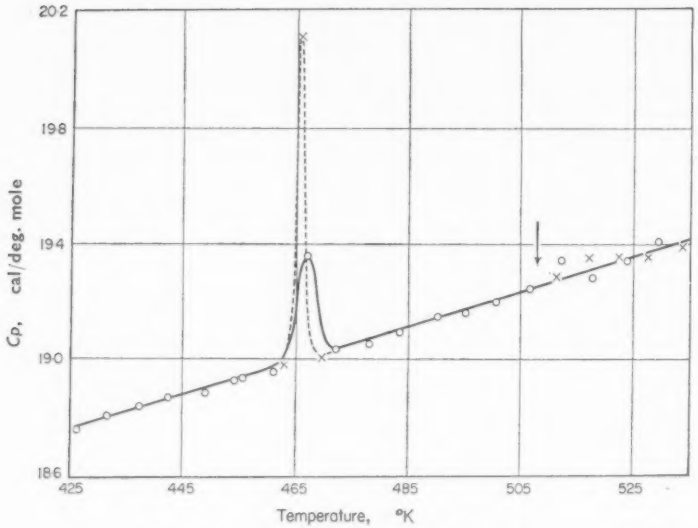


FIG. 3. Molar heat capacity of $MgNi_2$ ○–Series II, ×–Series III. Arrow shows Curie temperature reported by Voss[1].

Table 2. ΔC_p of formation of $MgNi_2$

Temperature (°K)	ΔC_p (cal. deg^{-1} mole^{-1})
10	−0·018
20	−0·010
30	+0·037
50	−0·11
70	−0·37
100	−0·51
150	−0·51
200	−0·52
250	−0·58
298·16	−0·80
300	−0·81
400	−1·23
500	−2·06
550	−2·90

Table 3. Entropy data for $MgNi_2$

Temperature (°K)	Third-Law entropy (e.u./mole)	Entropy of formation (e.u./mole)
273·16	19·68	—
298·16	21·20	−0·88
300	21·31	− ·90
400	26·52	−1·17
500	30·74	−1·54
555	32·76	—

quantities, the entropies of the elements were taken from the compilation by STULL and SINKE[10].

4. DISCUSSION OF RESULTS

The results of both the heat-capacity and susceptibility determinations conclusively demonstrate that $MgNi_2$ does not exhibit a Curie point at 235°C. The possibility that the ferromagnetism reported by Voss was induced by contact with and a slight absorption of, hydrogen was not disregarded. In this connection a sample of $MgNi_2$ was, as indicated above, sealed off under hydrogen, heat treated, and subjected to thermomagnetic analysis. The results, shown in Fig. 1, indicate that the susceptibility of the material is slightly reduced, rather than enhanced, by the hydrogen treatment.*

A more likely explanation is that the ferromagnetic behavior observed by Voss resulted from the

* The Curie point for the elemental nickel which is present as an impurity is slightly lowered. This and the lowered susceptibility suggest that the dissolved and adsorbed hydrogen is supplying a small number of electrons to the nickel *d*-orbitals.

substantial contamination of his materials by ferromagnetic impurities.

The susceptibility determinations indicate that $MgNi_2$ is a moderately strongly paramagnetic material. Its susceptibility is virtually the same at the two temperatures for which data are available, suggesting a temperature-independent susceptibility. This is the kind of behaviour expected from the collective-electron point of view if the material is exhibiting Pauli paramagnetism with a large density of states at the Fermi limit, as is the case with incomplete *d*-bands.

The results in Table 3 show that the compound is formed from its components with a reduction in entropy. This is in accordance with expectation in view of the high coördination number in the compound (average value is $13\frac{1}{3}$) and the consequent reduction of vibrational freedom (free volume) when it is formed.

VOL. 13 1960

The origin of the small thermal and magnetic effects observed at 194°C is a matter of considerable interest. One of the possibilities considered was that they were the result of a phase transformation at this temperature. In a recent publication, ELLIOTT has shown[11] that in certain intermetallic compounds of stoichiometry AB_2, transformations from one Laves-phase-type structure to another can occur with increasing temperature. The similarity of the structures implies a small heat of transformation in such cases, and indeed the heat effect discovered in $MgNi_2$ is only about 3 cal /mole. These considerations suggested that the effects at 194°C might be due to the transformation of $MgNi_2$ into either the $MgZn_2$ or $MgCu_2$-type structure. To establish whether this was indeed occurring, X-ray diffraction patterns were taken at 208 and 225°C. The patterns clearly indicated that the room-temperature structure of $MgNi_2$ was retained at these temperatures and excluded the possibility that a phase change was responsible for the phenomena observed.

Another possibility is that $MgNi_2$ is weakly antiferromagnetic, with its Néel point at 194°C. This seemed rather improbable in view of the smallness of the thermal effect, but an attempt was made to ascertain whether the susceptibility of $MgNi_2$ rises to a maximum in the vicinity of 194°C. It was, of course, necessary to eliminate by a Honda plot the effect of the elemental nickel present as impurity. As indicated earlier, in these experiments the maximum field attainable was insufficiently strong to permit a reliable extrapolation to infinite field, and so while the measurements failed to reveal a maximum in the susceptibility/temperature curve, the results were actually inconclusive.

Judging from the magnitude of the thermal and magnetic effects, it seems most likely that the anomalies originate with some type of irregularity to be found only at infrequent points in the $MgNi_2$ lattice. One likely possibility is the lattice defect which would occur when an occasional nickel atom replaces magnesium at a normal magnesium site in the lattice. This could occur either because of local fluctuations in concentration or as a result of the normal tendency of the system to increase its entropy by disordering so as to reduce its free energy. When magnesium is replaced by nickel, a 13-atom cluster of nickel atoms results. Exchange effects could be such as to cause these clusters to exhibit ferromagnetism at low temperatures. Qualitative considerations suggest that ferromagnetism in a system of this sort might be destroyed more abruptly with increasing temperature than in bulk elemental nickel. In so small an assembly the first disaligned spin could destabilize the entire system of aligned spins in the cluster, so that complete spin disorder would immediately ensue. As a result there would be an abrupt decrease in susceptibility over a very narrow temperature range. Furthermore, the energy of demagnetization would appear either as a latent heat or as an abnormally narrow λ point. These features are in keeping with the experimental observations at 194°C.

Although this seems at present to be the most likely of several possible explanations for the observed anomalies, it would be inappropriate to claim that the qualitative discussion above proves the correctness of the defect hypothesis. Confirmation of the thermal and magnetic effects of the nickel clusters, and indeed their existence, must await further, more critical, experimentation.

REFERENCES

1. VOSS G., *Z. anorg. Chem.* **57**, 61 (1908).
2. LAVES F. and WITTE H., *Metallwirtschaft* **14**, 645 (1935).
3. LAVES F. and WITTE H., *Mettallwirtschaft* **15**, 840 (1936).
4. CRAIG R. S. *et al.*, *J. Amer. Chem. Soc.* **76**, 238 (1954).
5. JOHNSTON W. V., STERRETT K. F., CRAIG R. S. and

WALLACE W. E., *J. Amer. Chem. Soc.* **79**, 3633 (1957).
6. HEDGECOCK F. T. and HUNT F., *Rev. Sci. Instrum.* **27**, 290 (1956).
7. HONDA K., *Ann. Phys. Lpz.* **32**, 1048 (1910).
8. SELWOOD P., *Magnetochemistry* p. 26. Interscience Publishers, New York (1956).
9. SCHÜBEL P. *Z., anorg. Chem.* **87**, 81 (1914).
10. STULL D. R. and SINKE G. C., *The Thermodynamic Properties of the Elements in their Standard States.* Dow Chemical Co., Midland, Mich. (1955).
11. ELLIOT R. P., *A Study of a Family of Laves-Type Intermediate Phases.* ASTIA Document No. AD-88033, Report from the Armour Research Foundation to Air Force Office of Scientific Research (1956).

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 13. pp. 221–234. Printed in Great Britain.

INFLUENCE DES CONDITIONS AUX LIMITES SUR LA DENSITE ELECTRONIQUE DANS UN RESEAU PERIODIQUE

G. LEMAN

Physique des Solides, Faculté des Sciences, Orsay (Seine-et-Oise) France

(*Received* 11 *July* 1959)

VOL. 13 1960

Résumé—Dans une structure de bandes, la densité électronique varie avec les conditions aux limites, au voisinage de ces limites. Cet effet est étudié en détail sur un modèle de Kronig et Penney. On montre que la contribution des bandes pleines décroît de façon exponentielle avec la distance r, en $\exp(-qr)$; une bande presque vide ou presque pleine donne un terme en $(\sin\Delta k_B r)/r$, où Δk_B est la différence en nombre d'onde entre le niveau de Fermi et la limite de bande la plus proche : ses électrons ou ses trous positifs se comportent comme s'ils étaient libres. Les valeurs exactes du paramètre q dépendent des bandes considérées et des conditions aux limites.

Abstract—In a band structure, the electronic density varies with the boundary conditions, near the boundaries. This effect has been studied in detail on a Kronig and Penney model. The contribution from full bands is shown to decrease exponentially with distance r, as $\exp(-qr)$; a nearly full or nearly empty band contributes a term $(\sin\Delta k_B r)/r$, where Δk_B is the difference in wave number between the Fermi level and the nearest band limit : its electrons or positive holes behave as though they were free. The exact values of the parameter q depend on the bands considered and on the boundary conditions.

1. INTRODUCTION

Si on détruit d'une façon quelconque dans un cristal parfait la périodicité du cristal, il s'ensuit une modification de la densité électronique qui peut s'étendre au-delà de la région perturbée dans la région où le potentiel est périodique. Cette perturbation à grande distance joue un rôle important dans certains problèmes physiques (alliages, magnétisme . . .). Nous nous proposons d'étudier cette perturbation dans la répartition électronique, autrement dit, l'influence de conditions aux limites quelconques sur la répartition électronique dans un cristal parfait.

Von Laue avait démontré, en étudiant les propriétés du rayonnement dans une cavité, que le nombre de degrés de liberté des photons était indépendant de la forme de la cavité aussi longtemps que l'on est à une distance des parois supérieure à plusieurs longueurs d'onde. Si l'on considère un gaz d'électrons libres enfermés dans une boîte de surface S, le théorème de Von Laue revient à dire que la densité électronique $\rho(E, \mathbf{r})$ au point $\mathbf{r}$, pour des électrons d'énergie moindre que E, est indépendante des conditions aux limites sur S pour tout point $\mathbf{r}$ de la boîte qui n'est pas

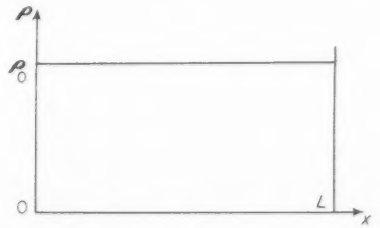


Fig. 1. Densité d'un gaz d'électrons libres avec les conditions aux limites de Born–von Karman.

trop près de S. A une dimension, par exemple, la densité électronique ρ_0 est constante dans le cas des conditions aux limites de Born–von Karman (Fig. 1); dans le cas des parois impénétrables

(Fig. 2), elle est telle que

$$\frac{\rho}{\rho_0} - 1 = -\frac{\sin 2kx}{2kx} \text{ avec } k^2 = 2E.$$

A des distances de la paroi supérieure à $\pi/2k$, la densité électronique est pratiquement indépendante des conditions aux limites.

Nous allons étudier en détail, dans cet article, la modification de la densité électronique dans un réseau unidimensionnel du type Kronig et Penney,

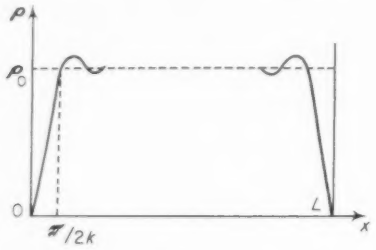


FIG. 2. Densité d'un gaz d'électrons libres dans le cas de deux parois impénétrables.

quand on introduit des conditions aux limites d'abord particulières (section 3) puis beaucoup plus générales (section 4). Auparavant, nous étudierons le problème de la détermination des fonctions d'onde électroniques (section 2).

2. DETERMINATION DES FONCTIONS D'ONDE

Considérons un réseau linéaire du type Kronig et Penney (Fig. 3). L'énergie des électrons est E, la longueur d'une maille a, les dimensions des

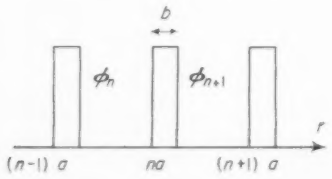


FIG. 3. Modèle de Kronig et Penney non perturbé.

barrières de potentiel b, V. Le problème n'est pratiquement pas modifié si l'on suppose les barrières de potentiel infiniment étroites; nous poserons $P = \lim_{\substack{b\to 0 \\ V\to\infty}} 2ab(V-E)$.

(a) *Fonctions d'onde dans le réseau non perturbé*

Nous partons des fonctions de Bloch d'énergie $E = \hbar^2k^2/2m$, soit :

$$\phi(k, r) = v(k_B, r) \exp(ik_B r) \text{ avec } v(k_B, r) \equiv v(k_B, r+a),$$

où k_B est un nombre réel à déterminer. Les conditions de continuité et de périodicité pour les fonctions ϕ et leurs dérivées, conduisent à la relation classique entre k et k_B

$$P\frac{\sin ka}{ka} + \cos ka = \cos k_B a. \tag{1}$$

Choisissons pour origine le milieu d'une maille (Fig. 4) et considérons une chaîne de longueur Na

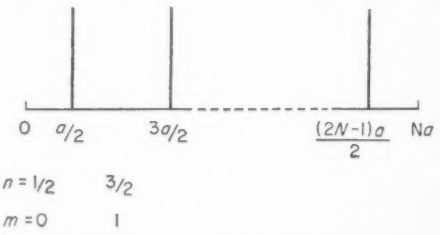


FIG. 4. Numérotation des mailles du réseau.

($N \gg 1$), à laquelle on applique les conditions aux limites de Born–von Karman. La fonction d'onde normalisée sera :

$$\phi(k_B, r) = v(k_B, r) \exp(ik_B r) \tag{2}$$

où

$$v(k_B, r) = a_0 \exp(-ik_B a/2) \times \times [\exp\{-i(k_B - k)r_n\} + + F\exp(-ika) \exp\{-i(k_B + k)r_n\}]. \tag{3}$$

Cette expression est valable pour $r_n = r - na$ compris entre $-a$ et 0. De plus :

$$F = \frac{\sin(k-k_B)a/2}{\sin(k+k_B)a/2}; \tag{4}$$

enfin ϕ est normalisée dans l'intervalle 0, Na si

$$a_0 = \left[Na\left(1 + F^2 + 2F\frac{\sin ka}{ka}\right)\right]^{-1/2}. \tag{5}$$

Rappelons la structure de bandes du modèle étudié afin de rendre plus clairs les développements

qui vont suivre. Avec les variables sans dimensions $u = ka$ et $x = k_B a$, la relation (1) s'écrit :

$$f(u) \equiv \frac{P \sin u}{u} + \cos u = \cos x. \qquad (6)$$

Si l'on représente $f(u)$ en fonction de u (Fig. 5), seules les parties de la courbe telles que $|f(u)| \leqslant 1$ doivent être retenues, puisque x est réel. L'axe

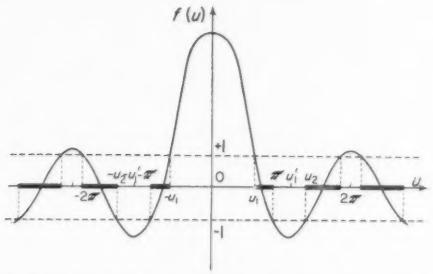


FIG. 5. Courbe $f(u)$. Les portions de l'axe des u en traits gras correspondent aux bandes d'énergies permises.

des u est ainsi divisé en portions comportant les valeurs permises de $u(u_n \leqslant |u| \leqslant n\pi)$ et d'autres interdites $(n\pi < |u| < u_{n+1})$. Les extremas de $f(u)$, obtenus pour des valeurs $u = 0, \ldots \pm u'_n$ telles que $n\pi < u'_n < u_{n+1}$ ont des ordonnées qui décroissent rapidement en module tout en restant supérieures à 1. Le cas limite des électrons libres (toutes les énergies sont permises) correspondrait à $P = 0$; celui des électrons liés correspond à P tendant vers l'infini (bandes permises infiniment étroites).

Puisque pour u donné, $x+2n\pi$ et x sont solutions, on peut, sans rien perdre de la généralité, faire correspondre conventionnellement à une valeur de u : $n\pi \leqslant u < (n+1)\pi$, les valeurs de x : $n\pi \leqslant |x| \leqslant (n+1)\pi$ (Fig. 6). En fait, avec les conditions aux limites de Born–von Karman, les branches de courbe $u(x)$ sont pseudo-continues, la densité d'états en x étant $\rho(x) = N/2\pi$, si l'on ne tient pas compte de la dégénérescence due au spin de l'électron.

(b) *Fonctions d'onde dans le réseau perturbé*

L'introduction dans un cristal parfait d'un défaut quelconque (impureté, surface libre) perturbe la répartition électronique. Autrement dit, la fonction d'onde électronique ne sera plus $\phi(k_B, r)$ mais une combinaison linéaire de fonctions de Bloch†

$$\psi(k_B, r) = Av(k_B, r) \exp(ik_B r) + \\ + Bv^*(k_B, r) \exp(-ik_B r). \qquad (7)$$

Les coefficients A et B sont déterminés par la normalisation des conditions aux limites de la fonction ψ. Nous ferons par la suite tendre la

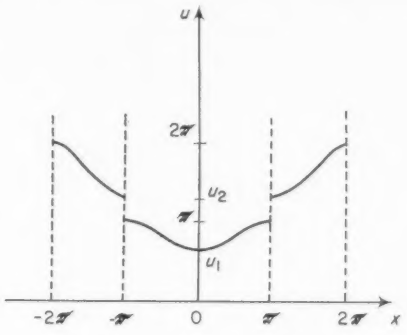


FIG. 6. Courbe $u(x)$.

taille Na du cristal vers l'infini, et nous nous intéresserons au voisinage d'une des limites seulement. Nous pouvons donc prendre des conditions aux limites arbitraires pour une des limites $r = 0$ et identiques pour l'autre $r = Na$.

$$\frac{1}{\psi}\left(\frac{\partial\psi(k_B, r)}{\partial r}\right)_{r=0} = \left(\frac{1}{\psi}\frac{\partial\psi}{\partial r}\right)_{r=Na} = h(k), \qquad (8)$$

où $h(k)$ est une fonction quelconque de l'énergie des électrons.

Nous allons montrer que (8) est équivalent aux trois conditions suivantes :

(i) $$\left(\frac{1}{\psi}\frac{\partial\psi(k_B, r)}{\partial r}\right)_{r=0} = h(k); \qquad (9)$$

(ii) $$\overline{|\psi(k_B, r)|^2} = \overline{|v(k_B, r)|^2} \text{ pour } r \text{ grand}, \qquad (10)$$

la moyenne étant faite sur une longueur grande par rapport à a;

(iii) Les ψ et les ϕ ont la même densité d'états par unité d'énergie.

† Nous supposerons que $\phi(k_B,r)$ et $\phi^*(k_B,r)$ sont linéairement indépendantes. En fait, il n'en est pas toujours nécessairement ainsi, en particulier au bord d'une bande.

VOL. 13 1960

(1) Il faut d'abord démontrer que la densité d'états $\rho(k_B)$ n'est pas modifiée par l'introduction des conditions (8). Celles-ci imposent $\exp(2ik_BNa) = 1$, donc $k_B = p(\pi/Na)$, p entier. Avec les conditions aux limites de Born–von Karman, on avait $k_B = 2p(\pi/Na)$, mais pour chaque valeur de $|k_B|$, il y avait dans ce cas deux états $\pm k_B$ de même énergie. L'introduction des conditions (8) lève en quelque sorte cette dégénérescence, et déplace l'ensemble des états vers les basses énergies. Le nombre total d'états tels que $E \leqslant E_M$ est légèrement modifié, mais non la densité d'états, qui reste égale à Na/π par unité de vecteur d'onde $|k_B|$.

(2) D'autre part, la condition (10) peut s'écrire :

$$|A|^2+|B|^2 = 1. \tag{11}$$

Or la condition de normalisation de $\psi(k_B, r)$ conduit à la même relation. En effet

$$\int_0^{Na} |\psi(k_B, r)|^2\, dr = |A|^2+|B|^2+$$

$$+\int_0^{Na} AB^*v^2(k_B, r)\exp(2ik_Br)\, dr+\text{c.c.} \tag{12}$$

A cause de sa périodicité, la fonction $v(k_B, r)$ peut être mise sous la forme $\exp(-ik_Ba/2)f(r_n)$ où $r_n = r - na$; l'intégrale figurant dans (12) s'écrit alors :

$$AB^*\left[N\int_{-a}^{0} f(r_n)\exp(2ik_Br_n)\, dr_n\right]\times$$

$$\times \sum_n \exp[i(2n-1)k_Ba].$$

Avec les notations choisies, n varie de $\frac{1}{2}$ à $(2N-1)/2$. Posons $2m = 2n - 1$; m représente alors le numéro de la maille compté à partir de l'origine, la maille contenant l'origine étant notée 0 (Fig. 4).

Puisque

$$k_Ba = p\frac{\pi}{N} \text{ et } \sum_{m=0}^{N-1} \exp[2imp(\pi/N)] = 0,$$

les deux intégrales de (11) sont nulles et les deux ensembles de conditions aux limites, (8) d'une part, (9) et (10) ou (11) d'autre part, sont équivalents et conduisent aux mêmes fonctions d'onde.†

Notre problème est finalement de déterminer, en fonction des conditions (9) et (11), comment la répartition électronique est modifiée par rapport à la densité périodique de Born–von Karman. Il faut évaluer l'intégrale :

$$I(r) = \int_{-k_{Bm}}^{k_{Bm}} \chi(k_B, r)\rho(k_B)\, dk_B, \tag{13}$$

où

$$\chi(k_B, r) \equiv |v(k_B, r)|^2 - |\psi(k_B, r)|^2$$
$$= 2Re'[-AB^*v^2(k_B, r)\exp(2ik_Br)], \tag{14}$$

l'intégration étant faite pour tous les électrons jusqu'au niveau de Fermi. Il serait compliqué d'évaluer exactement $I(r)$. Nous verrons que c'est une fonction décroissante en module et nous nous bornerons à évaluer la forme *asymptotique* de $I(r)$ pour r grand, d'abord dans un cas particulier de conditions aux limites, puis dans le cas général.

3. MODIFICATION DE LA DENSITE ELECTRONIQUE DUE A UNE BARRIERE INFINIE DE POTENTIEL

Nous supposerons qu'il existe à l'origine une *barrière infinie* de potentiel; au lieu de (9), nous aurons donc :

$$\psi(k_B, 0) = 0. \tag{15}$$

Par suite $A/B = -\exp(ika)$ et, compte tenu de (11), $|A| = |B| = \sqrt{2}/2$. La forme explicite de $\chi(k_B, r)$ est alors :

$$\frac{\chi(k_B, r)}{{a_0}^2} = \cos 2[kr-(k-k_B)ma]+$$

$$+F^2\cos 2[kr-(k+k_B)ma]+2F\cos 2k_Bma. \tag{16}$$

Cette expression est exacte mais compliquée; comme nous voulons chercher la forme asymptotique de $I(r)$, il est légitime de considérer son

† La forme explicite de $v(k_B,r)$ n'intervient pas dans la démonstration; celle-ci est donc valable pour un réseau unidimensionnel de potentiel quelconque : elle constitue une généralisation du théorème de von Laue aux réseaux périodiques unidimensionnels.

VOL. 13 1960

comportement moyen dans une maille. Nous considérerons donc

$$\chi(x,m)=\bar{\chi}(k_B,r)=\frac{1}{Na}\alpha(x)\cos 2mx$$

où

$$\alpha(x)=\frac{(1+F^2)(\sin u/u)+2F}{1+F^2+2F(\sin u/u)},$$

avec les variables sans dimensions $u=ka$, $x=k_Ba$. Puisque la densité d'états $\rho(x)$ est constante, égale à $N/2\pi$, la modification de la densité électronique sera donnée finalement par l'intégrale :

$$I(m)=\frac{1}{2\pi a}\int_{-x_M}^{x_M}\alpha(x)\cos 2mx\,dx. \qquad (17)$$

Tenant compte de (4) et (6), on peut mettre $\alpha(x)$ sous la forme

$$\alpha(u)=\frac{f(u)\cos u-(P+1)}{uf'(u)} \qquad (18)$$

$f'(u)$ désigant la dérivée de $f(u)$ par rapport à u.

Dans (17), l'intégration porte sur tous les électrons jusqu'au niveau de Fermi. Etant donné que les états d'énergies permises sont surtout concentrés aux limites de bandes, les cas les plus intéressants correspondent à plusieurs bandes pleines avec la dernière bande peu remplie, ou complètement pleine, ou presque pleine. Ayant résolu le problème pour une bande peu remplie puis pour une bande pleine, nous verrons que le cas d'une bande presque pleine se ramène aisément au premier en faisant intervenir les trous positifs.

(a) *Cas d'une bande peu remplie*

Raisonnons d'abord sur la première bande (Fig. 7). Il faut évaluer :

$$I(m)=\frac{1}{2\pi a}\int_{-x}^{x_M}\alpha(u)\cos 2mx\,dx$$

$$=\frac{1}{\pi a}\int_{0}^{x_M}\alpha(u)\cos 2mx\,dx.$$

Au voisinage de u_1, puisque $du/dx=0$ et $\alpha(u_1)=1$, on aura

$$\alpha(u)=1+C\frac{x^2}{2}\quad\text{avec}\quad C=\left(\frac{d^2\alpha}{dx^2}\right)_0.$$

Au premier ordre, l'intégrale $I(m)$ vaut

$$I(m)=\frac{1}{\pi a}\int_0^{x_M}\cos 2mx\,dx=-\frac{1}{\pi a}\frac{\sin 2mx_M}{m}. \qquad (19)$$

Elle décroît donc comme $(\sin 2k_Br)/r$ avec la distance, autrement dit les électrons de la première bande peu remplie se comportent comme des électrons libres.

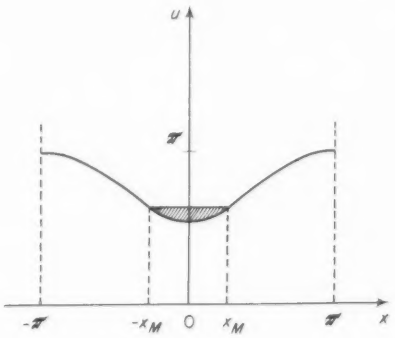


FIG. 7. Première bande peu remplie.

Si l'on tient compte du terme de 2ème ordre, la situation n'est pas modifiée puisque ce terme vaut :

$$\frac{C}{2\pi a}\int_0^{x_M}x^2\cos 2mx\,dx,$$

c'est-à-dire, en intégrant par parties,

$$\frac{C}{2\pi a}\left[x_M^2\frac{\sin 2mx_M}{2m}+x_M\frac{\cos 2mx_M}{2m^2}-\frac{\sin 2mx_M}{4m^3}\right].$$

Seul le premier terme est comparable avec le terme du 1er ordre, mais il reste négligeable si x_M est une petite fraction de π. Il varie d'ailleurs, avec la distance m, comme le terme du premier ordre dont il change seulement un peu l'amplitude des oscillations. Comme nous le verrons par la suite, le développement ci-dessus est valable pour

des électrons remplissant jusqu'à environ la moitié de la bande.

Bande quelconque. La situation est celle de la Fig. 8 et l'intégrale à évaluer est :

$$I(m) = \frac{1}{\pi a} \int_{(n-1)\pi}^{(n-1)\pi + x_M} \alpha(u) \cos 2mx\, dx.$$

En développant au voisinage de u_n et remarquant que $\alpha(u_n) = \pm 1$ suivant que n est pair ou impair, on aura encore au premier ordre :

$$|I(m)| \sim \frac{1}{\pi a} \frac{\sin 2mx_M}{m}. \qquad (20)$$

Par conséquent, à grande distance, les électrons de la bande se comporteront encore comme le même nombre d'électrons libres. La perturbation décroît avec la distance comme $(\sin 2\Delta k_B r)/r$ où Δk_B est la différence des nombres d'onde du niveau de Fermi et de la limite de bande.

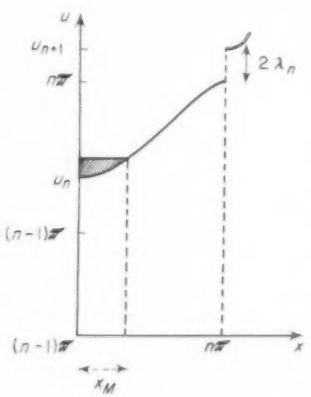


FIG. 8. $n^{\text{ième}}$ bande peu remplie.

(b) *Cas des bandes pleines*

Supposons maintenant que le niveau de Fermi se trouve dans une bande interdite. Nous allons montrer que les perturbations décroissent plus rapidement qu'une exponentielle.

(i) Pour la première bande pleine, l'intégrale exprimant la modification de la densité électronique sera :

$$I(m) = \frac{1}{2\pi a} \int_{-\pi}^{+\pi} \alpha(u) \cos 2mx\, dx$$

$$= \frac{1}{2\pi a} \int_{-\alpha}^{+\pi} \alpha(u) \exp(2imx)\, dx. \qquad (21)$$

Nous allons évaluer $I(m)$ en cherchant dans le plan complexe $z = x + iy$, un contour d'intégration pour la fonction $\alpha(z) \exp(2imz)$. Posons $w = u + iv$ avec la relation

$$f(w) \equiv P \frac{\sin w}{w} + \cos w = \cos z \qquad (22)$$

équivalente à (23) et (24).

$$\cos x \operatorname{ch} y = \frac{P}{u^2+v^2}(u \sin u \operatorname{ch} v + v \cos u \operatorname{sh} v) + \\ + \cos u \operatorname{ch} v \qquad (23)$$

$$\sin x \operatorname{sh} y = \frac{P}{u^2+v^2}(v \sin u \operatorname{ch} v - u \cos u \operatorname{sh} v) + \\ + \sin u \operatorname{sh} v. \qquad (24)$$

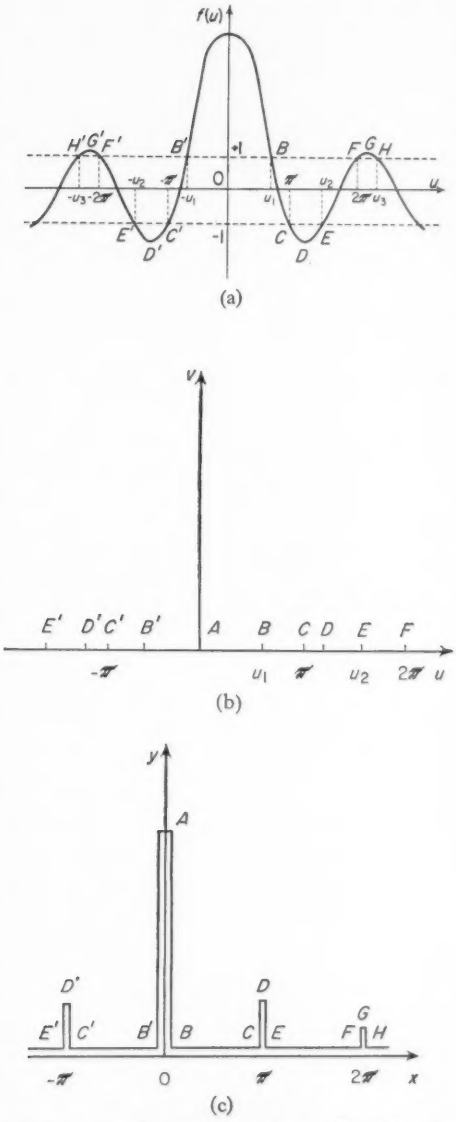
En posant

$$\alpha(w) = \frac{f(w) \cos w - (P+1)}{w f'(w)},$$

la fonction $\alpha(w) \exp(2imz)$ sera égale à $\alpha(u) \exp(2imx)$ sur l'axe réel ($y = 0$).

Le premier point à préciser est la correspondance entre les plans complexes w et z. Les résultats de cette étude (cf. Appendice) sont résumés sur les Fig. 9 a, b, c et 10 a et b. Quand u décrit l'axe réel du plan w de $-\infty$ à $+\infty$, le point correspondant du plan z décrit l'axe réel et les segments AB, AB', CD, $C'D'$. . . . Les points de l'axe réel $x'Ox$ correspondent aux bandes permises, les segments AB, CD . . . correspondent aux bandes interdites. En particulier les points A, D, D' . . . correspondent aux zéros de $f'(u)$. Pour que toute fonction uniforme de w soit aussi fonction uniforme de z, il faut couper le demi-plan supérieur ($y > 0$) par les segments AB, CD, $C'D'$. . . . Il est à noter cependant que le segment AB ne constitue pas

VOL. 13 1960



(a)

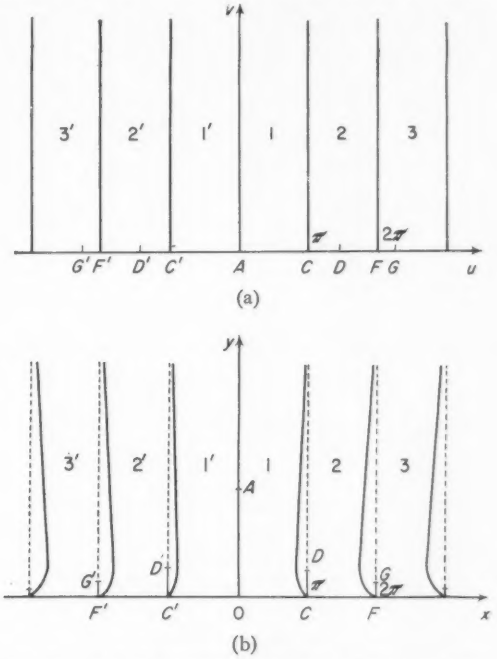
(b)

(c)

FIG. 9a, b, c. Représentation de l'axe réel du plan complexe w dans le plan complexe z.

une coupure pour toute fonction paire de w, telle que $\alpha(w)$. Enfin, la correspondance des demi plans $v > 0$ et $y > 0$ est représentée sur les Fig. 10 a et b.

Nous démontrons en appendice que la fonction $\alpha(z)$ n'a pas d'autres pôles que les points D, D', G, G'. . . . Il reste à choisir un contour d'intégration qui évite les segments CD, FG etc. . . .



(a)

(b)

FIG. 10a et b. Correspondance des demi-plans complexes w et z.

D'après les équations (23) et (24), toute fonction paire de w prend, en deux points symétriques par rapport à $0y$, des valeurs complexes conjuguées.

Considérons alors le contour d'intégration $C'CPP'$ (Fig. 11). Le long de ce contour

$$\int_{C'CPP'} \alpha(w) \exp(2imz)\, dz = 0$$

c'est-à-dire :

$$\int_{-\pi}^{+\pi} \alpha(u) \exp(2imx)\, dx +$$

$$+ \int_{\pi}^{-\pi} \alpha(w) \exp(2imx) \exp(-2mY)\, dx +$$

$$+ \int_{0}^{Y} \alpha(u)\, dy + \int_{Y}^{0} \alpha(u)\, dy = 0, \qquad (25)$$

soit

$$I + I' + I_1 + I_2 = 0.$$

L'intégrale I, au facteur $1/2\pi a$ près est celle que nous cherchons. Les intégrales I_1 et I_2 se compensent d'après la remarque faite précédemment,

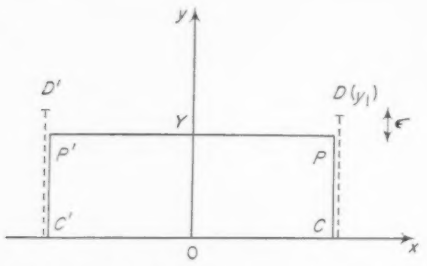


FIG. 11. Contour d'intégration pour la fonction $\alpha(z)e^{2imz}$ dans le cas de la première bande pleine.

puisque $\alpha(u)$ est paire en u. Reste l'intégrale I' qui peut s'écrire :

$$I' = \left[\int_{\pi}^{0} \alpha(w) \exp(2imx)\, dx + \right.$$

$$\left. + \int_{0}^{\pi} \alpha(w) \exp(2imx)\, dx \right] \exp(-2mY).$$

En faisant le changement x en $-x$ dans la deuxième intégrale et nous rappelant que, dans ce cas, $\alpha(w)$ devient $\alpha^*(w)$, on obtient :

$$I' = 2 \exp(-2mY) \int_{\pi}^{0} \text{Ré}\,[\alpha(w) \exp(2imx)]\, dx.$$

Appelons M la borne supérieure de $(2/\pi)|\alpha(w)|$ sur le segment $PP'(y = Y)$. On aura donc finalement :

$$|I(m)| < M \exp(-2mY). \qquad (26)$$

Nous avons choisi $Y < y_1$, y_1 correspondant au point D, c'est-à-dire au premier minimum de $f(u)$. Posons $Y = y_1 - \epsilon$, $\epsilon > 0$ aussi petit que l'on veut mais non nul. On pourra toujours trouver un nombre M indépendant de m, tel que

$$\frac{2}{\pi} |\alpha(w)|_{y=Y} < M,$$

car

$$|\alpha(w)| = \frac{|\cos z \,.\, \cos w - (P+1)|}{|wf'(w)|}.$$

Le numérateur est certainement inférieur à $\exp(2Y)$ et le dénominateur est certainement différent de 0 puisqu'il ne s'annule que pour $y = y_1$. Par conséquent, quand m croît, l'intégrale $I(m)$ tend vers 0 plus rapidement que l'exponentielle $M \exp(-2mY)$.

(ii) Ce qui précède concernait la 1eré bande pleine. Le raisonnement n'est pas modifié si l'on considère plusieurs bandes pleines. Ainsi pour la $n^{\text{ième}}$ bande, choisissons comme contour

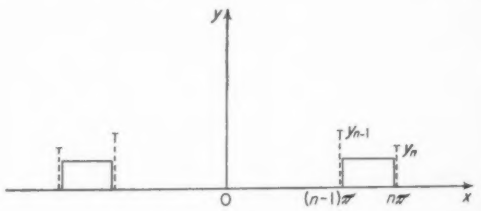


FIG. 12. Contour d'integration pour la fonction $\alpha(z)e^{2imz}$ dans le cas de la $n^{\text{ième}}$ bande pleine.

d'intégration le double contour de la Fig. 12. Il est clair que, si l'on pose :

$$I(m) = \int_{-n\pi}^{-(n+1)\pi} + \int_{(n+1)\pi}^{n\pi} \alpha(u) \exp(2imx)\, dx$$

$$= \tfrac{1}{2} \int_{(n-1)\pi}^{n\pi} \alpha(u) \cos 2mx\, dx,$$

on aura encore

$$|I(m)| \leqslant M \exp[-2m(y_n - \epsilon)], \qquad (27)$$

où $2M/\pi$ est la borne supérieure de $|\alpha(w)|$ pour $y = y_n - \epsilon$, $\epsilon > 0$ aussi petit que l'on veut mais non nul.

Par conséquent (1) le changement de densité électronique dû à l'introduction des conditions aux limites (10) et (15) dans le réseau de Kronig et Penney, décroît exponentiellement avec la distance à partir de l'origine quand on considère les électrons d'une bande pleine.

VOL. 13 1960

(2) S'il y a plusieurs bandes pleines, la contribution de la dernière bande s'étend plus loin que celle des autres bandes. Ceci ressort du fait que les valeurs de y_n décroissent rapidement comme le montre l'allure de la fonction $f(u)$ (Fig. 9a), dont les extremas $f(u'_n)$ satisfont à :

$$\operatorname{ch} y_n = |f(u'_n)|. \tag{28}$$

(3) Il est d'ailleurs facile de relier directement le coefficient de décroissance dans l'exponentielle $\exp[-2m(y_n-\epsilon)]$ à la largeur en $k(2\lambda_n = u_{n+1}-n\pi)$ de la bande interdite correspondante pour des bandes d'ordre élevé, c'est-à-dire des bandes interdites étroites (Fig. 8). On a simplement

$$y_n \simeq \lambda_n. \tag{29}$$

(iii) *Fonctions de Wannier.* KOHN[1] a récemment étudié les propriétés des fonctions de Wannier relatives à un modèle unidimensionnel où le potentiel $V(r)$ est périodique et symétrique. Il démontre que, pour chaque bande d'énergie, il existe une et une seule fonction de Wannier ayant les propriétés suivantes :

—elle est réelle;
—elle est symétrique ou antisymétrique;
—elle décroît exponentiellement :

$$W_n(r) \sim \exp(-h_n r).$$

n désigne le numéro de la bande considérée et h_n est la demi-largeur de la frange d'analyticité des fonctions de Bloch, c'est-à-dire exactement y_n dans le cas particulier que nous étudions. Il est intéressant de constater que ces résultats concordent avec les nôtres.

En effet, pour une bande quelconque (l'indice n a été supprimé pour ne pas surcharger l'écriture), la fonction de Wannier sera :

$$W(r) = \sqrt{(a/2\pi)} \int_{-\pi/a}^{\pi\alpha} \phi(k_B, r)\, dk_B, \tag{30}$$

où les $\phi(k_B, r)$ sont les fonctions de Bloch normalisées. Inversement, en désignant $W(r-la)$ par $W^l(r)$, on aura

$$\phi(k_B, r) = \sqrt{(a/2\pi)} \sum_1 W^l(r) \exp(ilk_B a). \tag{31}$$

Or la perturbation de la densité électronique pour une bande pleine peut s'écrire, en tenant compte de (14) et (31)

$$I(r) = \frac{1}{4\pi^2} \text{Ré} \sum_l \sum_{l'} W^l(r)\, W^{l'}(r) \times$$

$$\times \int_{-\pi}^{\pi} \exp(iu) \exp[i(l+l')x]\, dx. \tag{32}$$

Si nous admettons que les fonctions de Wannier sont réelles et se comportent comme $\exp(-mY)(Y \sim y_n \sim h_n)$, nous pouvons écrire :

$$I(m) = \frac{1}{4\pi^2} \sum_{l,l'} W^l(m)\, W^{l'}(m) \times$$

$$\times \int_{-\pi}^{+\pi} \cos[u+(l+l')x]\, dx, \tag{33}$$

où les fonctions de m sont les moyennes sur la maille m des fonctions correspondantes de r. Puisque u est fonction paire de x.

$$I(m) = \frac{1}{2\pi^2} \sum_{l,l'} W^l(m)\, W^{l'}(m) \int_0^{\pi} \cos u \,.\, \cos(l+l')x\, dx. \tag{34}$$

L'intégrale qui figure dans le second membre peut s'effectuer dans le plan complexe exactement comme l'intégrale (21). Elle donne un terme qui se comporte comme $\exp[-Y(l+l')]$. Puisque $W^l(m) \equiv W(m-1)$, on aura finalement:

$$I(m) \sim \sum_{l,l'} \exp[-Y(m-1)] \exp[-Y(m-l')] \times$$

$$\times \exp[-Y(l+l')]$$

c'est-à-dire $I(m)$ se comporte comme $\exp(-2mY)$: c'est bien le résultat (26).

(c) *Cas d'une bande presque pleine*

Nous sommes maintenant en mesure de traiter le cas d'électrons dont le niveau de Fermi se trouve près du sommet d'une bande (Fig. 13).

Soit en effet à évaluer

$$I(m) = \frac{1}{2\pi a} \int_{-x_M}^{x_M} \alpha(u) \cos 2mx \, dx. \qquad (35)$$

Nous raisonnerons sur la 1ère bande pour simplifier l'exposé; il est clair que le résultat vaudra pour n'importe quelle bande. Physiquement on

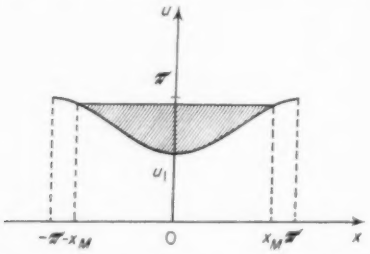


FIG. 13. Première bande presque pleine.

peut considérer que la contribution de cette bande est celle d'une bande pleine d'électrons, plus la contribution de trous correspondant aux valeurs $x_M < |x| < \pi$, avec la même densité d'états que les électrons de même nombre d'onde.

on aura pour $I''(m)$:

$$I''(m) = -\frac{1}{\pi a} \int_0^{\xi} \alpha(u) \cos 2m\xi \, d\xi$$

où u est relié à ξ par la relation

$$f(u) \equiv P\frac{\sin u}{u} + \cos u = -\cos \xi.$$

Développons au voisinage de $u = \pi$ et $\xi = 0$:

$$\alpha(u) = \alpha(\pi) + \frac{\xi^2}{2}\left(\frac{d^2\alpha}{d\xi^2}\right)_0 = 1 + D\frac{\xi^2}{2}. \qquad (37)$$

Au premier ordre, on aura :

$$I''(m) = \frac{1}{\pi a}\frac{\sin 2m\xi_M}{m} = \frac{1}{\pi a}\frac{\sin 2mx_M}{m}. \qquad (38)$$

Les termes du 2nd ordre seront en

$$\xi_M{}^2 \frac{\sin 2m\xi_M}{2m} + 0\left(\frac{1}{m^2}\right)$$

donc négligeables si ξ_M est une petite fraction de π.

Puisque l'intégrale $I'(m)$ dans (36) correspond

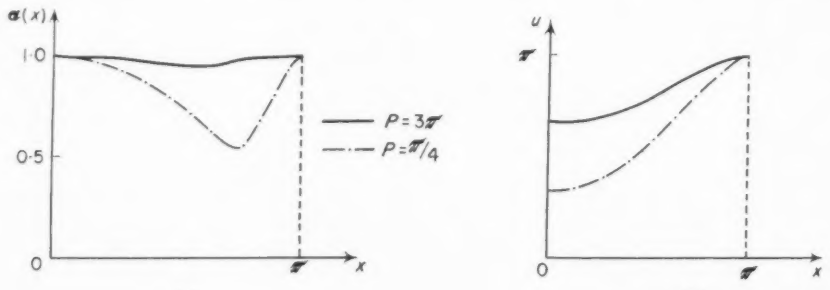


FIG. 14 a et b. Courbes $\alpha(x)$ et $u(x)$ pour deux valeurs particulières de P.

Mathématiquement, il suffit de faire le changement de variable $\xi = \pi - x$. Puisque.

$$I(m) = I'(m) + I''(m) = \frac{1}{\pi a}\int_0^{\pi} \alpha(u) \cos 2mx \, dx -$$

$$-\frac{1}{\pi a}\int_{x_M}^{\pi} \alpha(u) \cos 2mx \, dx \qquad (36)$$

à une bande pleine elle décroît exponentiellement avec la distance et sa contribution est négligeable aux grandes distances m, par rapport à $I''(m)$. Par conséquent la perturbation de la densité électronique, dans le cas d'une bande presque pleine, décroît avec la distance comme

$$\frac{\sin 2\Delta k_B r}{r} \text{ où } \Delta k_B = \frac{\pi}{a} - k_{BM}$$

les trous se comportent comme des trous libres.

VOL. 13 1960

Validité de l'approximation. On peut se demander jusqu'à quel point le développement en puissances de x ou $\pi - x$ qui intervient dans 3(a) et 3(c) est valable. Pour s'en faire une idée, on a évalué numériquement la fonction $\alpha(x)$ dans deux cas très différents : $P = 3\pi$ (première bande étroite) et $P = \pi/4$ (première bande assez large) (Fig. 14 a et b).

(1) Pour une bande étroite, $\alpha(x)$ diffère peu de l'unité (il est compris entre 0,96 et 1 pour $P = 3\pi$). La courbure de $\alpha(x)$ sera toujours faible et on aura avec une très bonne approximation

$$I(m) \simeq \frac{-1}{a} \frac{\sin 2mx_M}{m}$$

avec x_M quelconque entre 0 et π.

(2) Dans le cas général, $\alpha(x)$ présente un minimum assez marqué pour une valeur x_m d'autant plus proche de π que P est faible, c'est-à-dire que la bande est large (Fig. 15 a et b). La 1ère partie de

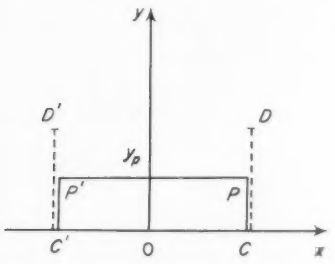


FIG. 15. Contour d'intégration de la fonction $1-g(w)/1+g(w)\ \alpha(z)e^{2imz}$ dans le cas de la première bande pleine.

la courbe ($0 < x_M < x_m$) présente une courbure presque constante et par conséquent, il suffira dans le développement de $\alpha(x)$ en puissances de $x(3a)$ de prendre les termes jusqu'au 2e ordre inclus, pour avoir une bonne approximation, avec des électrons remplissant jusqu'à la moitié de la bande ($x_M \leqslant \pi/2$).

(3) Par contre, la 2ème partie de la courbe $\alpha(x)$ ($x_m < x_M < \pi$) présente une forte courbure au voisinage de π. Le développement en puissances de π-x ne sera donc valable que si le niveau de Fermi est proche de la limite de la bande et les termes du second ordre joueront un rôle important.

4. CONDITIONS AUX LIMITES GENERALES

Il nous faut maintenant considérer le cas général où la fonction d'onde $\psi(k_B,r)$ est déterminée par les conditions (9) et (11).

La fonction $h(k)$ est réelle ou complexe. Ainsi, par exemple, avec les conditions aux limites de Born–von Karman,

$$h(k) = \frac{ik(1-F)}{1+F},$$

tandis que pour une barrière de potentiel finie à l'origine ($V(r) = V_0$ pour $r \leqslant 0$) $h(k) = \sqrt{(2V_0 - k^2)}$. Introduisons le déphasage η défini par

$$\frac{A}{B} = \exp[i(\eta + ka)] \tag{39}$$

et posons

$$c = 2|A|\,|B|. \tag{40}$$

On montre facilement que la modification de la densité électronique est alors donnée par :

$$I(m) = \frac{-1}{2\pi a} \int_{-x_M}^{x_M} c \cos\eta\alpha(u) \cos 2mx\, dx. \tag{41}$$

D'autre part, en explicitant (9) et tenant compte de (12), on obtient

$$c \cos\eta = \frac{1-g(u)}{1+g(u)} \tag{42}$$

à condition de poser

$$g(u) = \left| \frac{h(u/a)}{u/a} \frac{1+F}{1-F} \right|^2. \tag{43}$$

Dans le cas particulier où $h(k)$ est réel, on a $c = 1$, c'est-à-dire $|A| = |B| = \sqrt{(2)/2}$, donc $\eta = -2$ arc tg $\sqrt{[g(u)]}$. Par ailleurs, pour les valeurs de l'énergie correspondant aux limites de bande ($u = u_n$ ou $n\pi$), on a toujours $|c \cos\eta| = 1$.

L'évaluation de $I(m)$ dans le cas d'une bande peu remplie se fait alors exactement de la même manière que dans la section 3. Ici encore, on trouve que les électrons se comportent à grande distance, comme des électrons libres : la perturbation de la densité électronique décroît comme $(\sin 2\Delta k_B r)/r$ où Δk_B est la différence de nombre

d'onde entre le niveau de Fermi et la limite de bande la plus proche.

Quand le niveau de Fermi se trouve dans une bande interdite, l'intégrale $I(m)$ est de la forme :

$$I(m) = \frac{1}{2\pi a} \int_{-\pi}^{\pi} \frac{1-g(u)}{1+g(u)} \alpha(u) \exp(2imx)\, dx.$$

La méthode d'intégration dans le plan complexe $z = x+iy$ peut encore être utilisée à condition de poser

$$D(u) = a\,|h(u/a)|$$

et

$$g(w) = \left(\frac{D(w)\,\mathrm{tg}(w/2)}{w\,\mathrm{tg}(z/2)} \right)^2.$$

La fonction $1-g(w)/1+g(w)$ peut évidemment introduire de nouveaux pôles mais certainement pas sur l'axe réel du plan w c'est-à-dire sur l'axe $x'0x$ et sur les segments CD, $C'D'$ ... Il suffit alors de prendre comme contour d'intégration (Fig. 15) un rectangle $C'CPP'$ laissant à l'extérieur tous les pôles de $1-g(w)/1+g(w)$. Soit y_p la hauteur de ce rectangle; on aura encore

$$|I(m)| < M \exp(2my_p).$$

La perturbation de la densité électronique, due à l'introduction de conditons aux limites quelconques, décroît donc exponentiellement avec la distance, quand le niveau de Fermi se trouve dans une bande interdite. *Mais nous ne pouvons pas fixer d'une façon générale la valeur exacte de cette exponentielle.*

Enfin le traitement d'une bande presque pleine est absolument semblable à celui qui a été décrit précédemment; ici encore, la perturbation de la densité électronique décroît avec la distance comme $(\sin 2\Delta k_B r)/r$, Δk_B étant la différence entre les nombres d'onde du niveau de Fermi et de la limite de bande : les trous positifs se comportent comme des trous libres.

5. CONCLUSIONS

L'introduction d'un défaut localisé quelconque dans un réseau périodique du type Kronig et Penney perturbe la répartition électronique. Cette perturbation s'étend à plus ou moins grande distance suivant que l'on considère les électrons d'une bande pleine ou d'une bande peu remplie ou presque pleine.

(1) Quand le niveau de Fermi se trouve dans une bande d'énergie permise et près d'une limite de bande, nous avons montré que les électrons de bas de bande ou les trous positifs de la bande presque pleine se comportent comme un gaz d'électrons ou de trous libres de même densité moyenne. La perturbation de la densité électronique varie à grande distance r comme $(\sin 2\Delta k_B r)/r$, Δk_B étant la différence entre les nombres d'onde du niveau de Fermi et de la limite de bande la plus proche. Pour les gaz denses (Δk_B grand), cette expression est meilleure pour des électrons que pour des trous, surtout dans les bandes assez larges.

(2) Les contributions des bandes pleines décroissent plus vite avec r, et de façon exponentielle, en $\exp(-qr)$. Elles sont donc négligeables devant les premières à grandes distances. Quand on remplace les conditions périodiques de Born–von Karman par une barrière de potentiel impénétrable la charge repoussée localement dans une bande pleine a le même coefficient q que les fonctions de Wannier de la bande; q croît avec la largeur des bandes interdites voisines; pour des électrons presque libres, $aq \simeq \Delta k$, si a est le paramètre du réseau, Δk la largeur en k de la bande interdite au sommet de la bande et $\hbar^2k^2/2m$, est l'énergie. Le coefficient q peut par contre être plus petit pour d'autres conditions aux limites.

Ces conclusions sont en accord qualitatif avec celles de BLOEMBERGEN et ROWLAND[2] pour un réseau à trois dimensions perturbé au premier ordre. Le modèle de Kronig et Penney est trop théorique pour que des comparaisons quantitatives soient possibles avec des données expérimentales; il faudrait pour cela étudier l'influence des conditions aux limites sur la densité électronique dans un réseau à trois dimensions, en utilisant, par exemple, une méthode de perturbations. Néanmoins si les résultats obtenus restent essentiellement les mêmes pour un réseau quelconque, ils peuvent donner des renseignements sur le problème des impuretés dans les semi-métaux (As, Bi ...), où la bande de valence est presque pleine et la bande de conductibilité presque vide : quand on y introduit une impureté elle perturbera la densité électronique à grande distance; en particulier deux impuretés pourront interagir à

grande distance. De même, le magnétisme des métaux de transition pourrait être relié de cette façon à leur structure de bande : Si la bande $3d$ est presque pleine (Fe,Co,Ni), l'aimantation d'un atome, considérée comme une perturbation locale, agira à grande distance et produira un couplage ferromagnétique sur les atomes voisins. Si au contraire le niveau de Fermi se trouve vers le milieu de la bande (Cr, Mn), l'aimantation d'un atome produit autour de lui une polarisation de spin qui oscille plus rapidement dans l'espace et peut tendre à aligner ses voisins antiparallèlement à lui-même.

Remerciement—Je tiens à exprimer ici ma profonde gratitude à Monsieur le Professeur J. FRIEDEL qui m'a constamment guidé au cours de ce travail.

VOL. 13 1960

APPENDICE

Etude de la correspondance entre les plans complexes

$$w = u+iv \text{ et } z = x+iy$$

Les équations de la correspondance sont :

$$\cos x \operatorname{ch} y = \frac{P}{u^2+v^2}(u \sin u \operatorname{ch} v + v \cos u \operatorname{sh} v) + \cos u \operatorname{ch} v \qquad (A1)$$

$$\sin x \operatorname{sh} y = \frac{P}{u^2+v^2}(v \sin u \operatorname{ch} v - u \cos u \operatorname{sh} v) + \sin u \operatorname{sh} v \qquad (A2)$$

Les équations (A1) et (A2) ne changent pas si l'on remplace w par $-w$ et z par $-z$. Nous n'aurons donc à étudier que la correspondance entre les demi-plans supérieurs $y \geqslant 0$ et $v \geqslant 0$ puisque de toute façon l'on prendra le contour d'intégration dans ces demi-plans.

(1) *Axe réel du plan* w(v = 0)

Les équations (A1) et (A2) s'écrivent dans ce cas :

$$\cos x \operatorname{ch} y = \frac{P \sin u}{u} + \cos u \qquad (A3)$$

$$\sin x \operatorname{sh} y = 0. \qquad (A4)$$

Représentons la fonction $f(u)$, u variant de $-\infty$ à $+\infty$, en désignant par A, B, B', C, C', D, D' etc. ... les points d'ordonnée extremum ou d'ordonnée égale à ± 1 (Fig. 9a).

Par continuité, en partant de $u = 0$ (point A sur les trois figures 9a, b et c), l'examen des équations (A1) et (A2) montre qu'à l'axe réel du plan w, on peut faire correspondre l'axe réel du plan z, augmenté des segments AB, CD ..., parcourus deux fois. Ces segments seront des coupures dans le demi-plan $y>0$ car sur les bords droit et gauche de CD, par exemple, u n'a pas la même valeur. Cependant, le segment AB ne sera pas une coupure pour les fonctions paires de w, telles que $f(w)$ et $\alpha(w)$.

(2) *Demi-plan* $v>0$

Faisons maintenant varier v de 0 à $+\infty$, et divisons le demi-plan $v>0$ en zones 1, 1′, 2, 2′ ... (Fig. 10a). Le choix de détermination que nous venons de faire pour u, impose par continuité, de faire correspondre à ces zones du plan w, les zones 1, 1′ ... du plan z (Fig. 10b). Les courbes limites de ces zones sont déterminées par les équations (A5) et (A6) (où l'on a pris $u = n$) :

$$\cos x \operatorname{ch} y = \pm \frac{P}{n^2\pi^2+v^2} v \operatorname{sh} v \pm \operatorname{ch} v; \qquad (A5)$$

$$\sin x \operatorname{sh} y = \mp \frac{P}{n^2\pi^2+v^2} \operatorname{sh} v \qquad (A6)$$

Il faut prendre le signe + dans (A5), le signe − dans (A6), si n est pair et inversement.

Précisions la forme de ces courbes pour n positif (pour n négatif, il suffira de prendre les symétriques par rapport à $0y$). Suivant que n est pair ou impair, on aura $\cos x \gtrless 0$ et $\sin x \gtrless 0$; en tenant compte du fait que x tend vers $n\pi$ quand v et y tendent vers 0 (puisque $u = n\pi$), on voit que la courbe correspondante à la droite $u = n\pi$ sera nécessairement dans la zone comprise entre $x = n\pi$ et $x = (n-\frac{1}{2})\pi$.

Pour v tendant vers ∞, on aura :

$$\cos x \operatorname{ch} y \simeq \pm \operatorname{ch} v; \qquad (A7)$$

$$\sin x \operatorname{sh} y \simeq \mp P n\pi \frac{\operatorname{sh} v}{v^2}. \qquad (A8)$$

On voit que x tend vers $n\pi$ et y vers ∞, l'écart entre la courbe et son asymptote tendant vers 0 comme $Pn\pi/y^2$, quand v et y tendent vers ∞.

Pour obtenir enfin le comportement de la courbe au voisinage de l'axe réel, il suffit de développer les équations (A5) et (A6) en puissances de v. On montre ainsi que la courbe correspondant à $u = n\pi$ admet une tangente à 45° pour y et v tendant vers 0.

On notera que pour $u = 0$, l'équation (A6) se réduit à :

$$\sin x \operatorname{sh} y = 0.$$

L'axe imaginaire du demi-plan $v>0$, a donc pour correspondant la demi-droite Ay ($x = 0$, $y>y_0$) du demi plan $y>0$.

Finalement, la correspondance entre les plans w et z est celle que nous avons représentée Fig. 10 a et b.

Pôles de la fonction $\alpha(w)$

La fonction :

$$\alpha(w) = \frac{f(w)\cos w - (P+1)}{wf'(w)}$$

n'a de pôles que sur l'axe réel ($v = 0$) du plan complexe w.

En effet, les pôles de $\alpha(w)$ correspondant aux valeurs de w qui annulent $f'(w)$. Donc :

$$f'(w) = \frac{P\cos w}{w} - \frac{P\sin w}{w^2} - \sin w = 0 \quad \text{(A9)}$$

ce qui peut s'écrire :

$$\operatorname{tg} w = \frac{Pw}{P+w^2} \quad \text{(A10)}$$

soit

$$\frac{\operatorname{tg}^2 u + \operatorname{th}^2 v}{(1 - i\operatorname{tg} u \operatorname{th} v)(\operatorname{tg} u - i \operatorname{th} v)}$$

$$= \frac{P(u^2+v^2)}{(u-iv)(P+u^2-v^2+2iuv)}.$$

En inversant et identifiant parties réelle et imaginaire, on obtiendra :

$$\begin{aligned} \frac{\operatorname{tg} u}{\operatorname{ch}^2 v} \quad \frac{1}{\operatorname{tg}^2 u + \operatorname{th}^2 v} &= u \frac{P+u^2+v^2}{P(u^2+v^2)} \\ \frac{\operatorname{th} v}{\cos^2 u} \quad \frac{1}{\operatorname{tg}^2 u + \operatorname{th}^2 v} &= v \frac{P-u^2-v^2}{P(u^2+v^2)}. \end{aligned} \quad \text{(A11)}$$

Si $v \neq 0$, on peut faire le rapport de ces expressions et écrire

$$\frac{\sin 2u/2u}{\operatorname{sh} 2v/2v} = \frac{P+u^2+v^2}{P-u^2-v^2}.$$

Le module du second membre est supérieur à 1 et celui du premier inférieur à 1 : il faut donc que $v = 0$. Les 0 de $f'(w)$ sont les 0 de $f'(u)$.

Par conséquent, les pôles de $\alpha(w)$ seront dans le plan complexe z, les points D, D', G, G' etc. ... (Fig. 9c). Mais le point A, c'est-à-dire le point correspondant à $u = 0$, n'est pas un pôle pour $\alpha(u)$ car

$$\lim_{u \to 0} \alpha(u) = \frac{2P+3}{P+3}.$$

BIBLIOGRAPHIE

1. Kohn W., *Phys. Rev.* **115,** 809 (1959).
2. Bloembergen N. et Rowland T. J., *Phys. Rev.* **97,** 1679 (1955).

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 13. pp. 235–243. Printed in Great Britain.

ETUDE EXPERIMENTALE DE LA DIFFUSION CRITIQUE DES NEUTRONS DANS LE FER

M. ERICSON* et B. JACROT†

(*Received* 20 *June* 1959)

VOL. 13 1960

Abstract—The cross section for critical scattering of slow neutrons by iron near the Curie temperature is studied.

The parameters κ_1, μ_1, Λ introduced in the VAN HOVE's theory are measured for various temperatures.

For instance, at a temperature of 826°C ($T-T_c = 56°$) it is found:

$$\kappa_1^2 = 1{\cdot}66 \times 10^{-2}(\text{Å})^{-2}, \frac{2m\Lambda}{\hbar} = 3{\cdot}5.$$

The results for κ_1 and μ_1 disagree with the results of other authors. They are consistent with magnetic data.

The results are compared with different theoretical estimations. The best fit is obtained with a model in which the second neighbours of an iron atom are magnetically active and in which the magnetic electrons are described by the Heisenberg model.

Résumé—On étudie la section efficace de diffusion critique des neutrons par le fer au voisinage de la température de Curie.

On mesure les paramètres introduits dans la théorie de VAN HOVE, κ_1, μ_1 et Λ, à diverses températures.

On trouve par exemple à une température de 826°C ($T-T_c = 56°$)

$$\kappa_1^2 = 1{,}66 \cdot 10^{-2}(\text{Å})^{-2}, \frac{2m\Lambda}{\hbar} = 3{,}5.$$

Les résultats pour κ_1 et μ_1 sont en désaccord avec ceux d'autres auteurs. Ils sont compatibles avec les données magnétiques.

On compare les résultats à différentes prévisions théoriques. Parmi les modèles essayés, celui qui donne le meilleur accord est un modèle dans lequel les seconds voisins d'un atome de fer sont magnétiquement actifs et où les électrons magnétiques sont décrits par le modèle de Heisenberg.

1. INTRODUCTION

LES fluctuations d'aimantation qui existent dans un corps ferromagnétique au voisinage de la température critique sont responsables d'une diffusion magnétique anormale des neutrons, découverte expérimentalement dans le fer par PALEVSKY[1] et SQUIRES[2] et expliquée par VAN HOVE[3, 4].

Soit M le moment magnétique résultant d'une région du cristal ferromagnétique contenant un grand nombre N d'atomes. On a:

$$\langle M^2 \rangle - \langle M \rangle^2 = 3Nk_BT\chi$$

χ est la susceptibilité par atome, T la température de l'échantillon, k_B la constante de Boltzman.

A la température critique la susceptibilité est infinie, les fluctuations sont également infinies.

Ces fluctuations peuvent se décrire en termes de corrélations; on définit une fonction de corrélation dépendant du temps.

$$\gamma_R(t) = \langle \boldsymbol{S}_0(0)\boldsymbol{S}_R(t) \rangle ,$$

valeur moyenne du produit scalaire d'un spin de

* Centre National de la Recherche Scientifique.

† Centre d'Études Nucléaires de Saclay, Gif-sur-Yvette (S-et-O), France.

référence et d'un spin situé à la distance $\boldsymbol{R}$ du premier et considéré à l'instant t. La fonction de corrélation est proportionnelle à la probabilité de trouver au temps t et à l'endroit défini par le vecteur $\boldsymbol{R}$ un spin aligné avec le spin de référence.

VAN HOVE a montré que la section efficace de diffusion des neutrons, par unité d'angle solide et unité d'énergie des neutrons diffusés, est proportionnelle à la transformée de Fourier à 4 dimensions de la fonction de corrélation

$$\frac{d^2\sigma}{d\Omega d\epsilon} = \left(\frac{ge^2}{m_0c^2}\right)^2 \frac{1}{3\pi\hbar}\frac{k'}{k}|f(K)|^2 \times$$

$$\times \sum_R \int \exp\left[i(\boldsymbol{KR}-\omega t)\right]\gamma_R(t)\,dt$$

$g = 1{,}91$	moment magnétique du neutron, mesuré en magnétons nucléaires,
e^2/m_0c^2	rayon classique de l'électron,
e	énergie des neutrons diffusés,
$\boldsymbol{k}'$	vecteur d'onde des neutrons diffusés,
$\boldsymbol{k}$	vecteur d'onde des neutrons incidents
$\boldsymbol{K} = \boldsymbol{k}'-\boldsymbol{k}$	vecteur diffusion, $K = \lvert\boldsymbol{K}\rvert$
$\hbar\omega$	changement d'énergie des neutrons lors de la diffusion,
$f(K)$	facteur de forme, égal à 1 pour $K=0$.

La somme Σ_R est étendue à tous les atomes magnétiques du réseau.

L'étude expérimentale de la section efficace permet d'obtenir des renseignements sur la fonction de corrélation. Nous avons fait cette étude dans le cas du fer.

Au voisinage de la température de Curie, le temps de relaxation des fluctuations est assez grand pour qu'on puisse, dans une première étude, négliger l'évolution dans le temps de la fontion de corrélation. La section efficace par unité d'angle solide $d\sigma/d\Omega$ ne dépend que de la variation dans l'espace de la fonction de corrélation.

Notre étude comprend deux séries d'expériences : dans une première série, les neutrons incidents ont des énergies réparties dans un spectre large; on obtient des renseignements uniquement sur l'étendue spatiale de la fonction $\gamma(\boldsymbol{R})$; dans une deuxième série d'expériences, les neutrons incidents sont approximativement monocinétiques; on peut alors étudier les changements d'énergie des neutrons lors de la diffusion et obtenir ainsi des renseignements sur la cinétique des fluctuations.

2. ÉTUDE DE LA DÉPENDANCE SPATIALE DE LA FONCTION DE CORRELATION

A la température de Curie, si l'on fixe la direction d'un spin, des spins éloignés du premier ont tendance à s'aligner sur ce spin de référence; la fonction de corrélation des spins décroît alors très lentement vers 0 quand la distance augmente. Au fur et à mesure que la température s'élève, un spin n'influence que des voisins de plus en plus proches, la fonction de corrélation décroît de plus en plus rapidement vers 0. La fonction $\gamma(R)$ a un profil en "toile de tente". A l'origine, elle a pour valeur $S(S+1)$, corrélation d'un spin avec lui-même. VAN HOVE a montré que pour des distances assez grandes on peut la caractériser au moyen de 2 paramètres, κ_1 et μ_1, permettant de calculer respectivement la portée et l'intensité des corrélations.

L'expression asymptotique de la fonction de corrélation est la suivante :

$$\gamma(R) = \frac{v_0 S(S+1)}{4\pi\mu_1^2}\,\frac{\exp(-\kappa_1 R)}{R}$$

v_0 est le volume atomique, S le spin.

Si l'on remplace la fonction de corrélation par son expression asymptotique dans l'expression de la section efficace, on obtient :

$$\frac{d\sigma}{d\Omega} = \frac{2}{3}\left(\frac{ge^2}{m_0c^2}\right)^2 \frac{S(S+1)}{\mu_1^2}\,\frac{1}{\kappa_1^2+K^2}$$

Cette formule est valable pour de petits angles de diffusion tels que $K \ll 2\pi\tau_{\min}$, où $\tau_{\min}$ est la longueur du plus petit vecteur du réseau réciproque. Cette condition était réalisée dans nos expériences où la valeur maximum de K était de 0,17 $(\text{Å})^{-1}$ alors que $2\pi\tau_{\min}$ dans le fer est égal à $3{,}1(\text{Å})^{-1}$.

κ_1 et μ_1 sont reliés par la relation :

$$(\kappa_1\,\mu_1)^2 = \frac{\chi_0}{\chi}$$

dans laquelle χ_0 est la susceptibilité d'un atome isolé et χ la susceptibilité effective par atome.

A la température critique, la susceptibilité est infinie, κ_1 est nul, les corrélations ont une portée infinie.

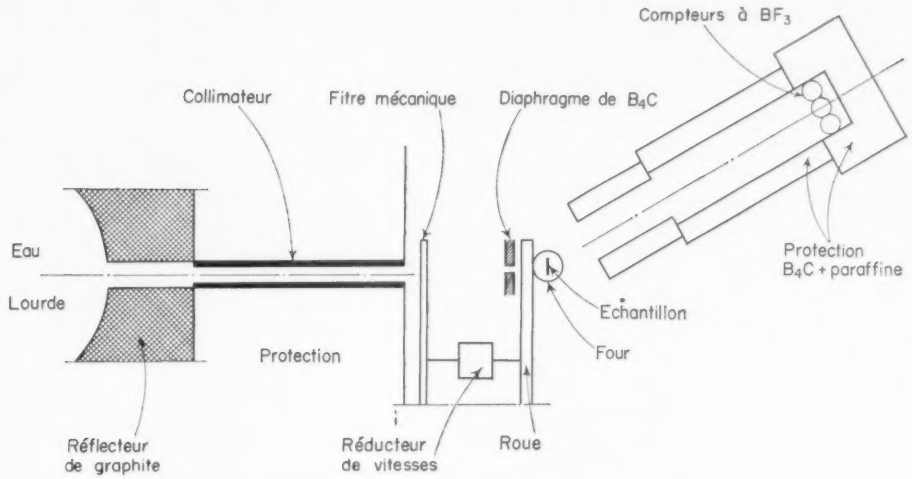


FIG. 1. Schéma général du dispositif expérimental.

L'étude de la section efficace différentielle $d\sigma/d\Omega$ permet la détermination des paramètres κ_1 et μ_1 qui interviennent dans son expression. On peut étudier cette section efficace soit en mesurant la distribution angulaire de neutrons de longueur d'onde donnée (on fait alors varier K par l'intermédiaire de θ), soit en mesurant la variation, en fonction de la longueur d'onde, de la section efficace de diffusion à un angle donné; on fait alors varier K par l'intermédiaire de λ. Nous avons utilisé cette deuxième méthode. Le principe des expériences est le suivant: un faisceau de neutrons tombe sur un échantillon de fer chauffé. On étudie par une méthode de temps de vol le spectre des neutrons diffusés à un angle de quelques degrés (Fig. 1).

(a) *Résultats à la température de Curie*

A cette température $\kappa_1 = 0$. La section efficace a alors pour expression:

$$\frac{d\sigma}{d\Omega} = \frac{2}{3}\left(\frac{ge^2}{m_0c^2}\right)^2 \frac{S(S+1)}{\mu_1^2}\,\frac{1}{K^2}$$

$$= \frac{2}{3}\left(\frac{ge^2}{m_0c^2}\right)^2 \frac{S(S+1)}{\mu_1^2}\,\frac{\lambda^2}{4\pi^2\theta^2}$$

(si θ est petit).

Elle est proportionnelle au carré de la longueur d'onde. Nous l'avons vérifié en comparant le spectre $I_{T_C}(\lambda)$ des neutrons diffusés à un angle de 3 degrés par le fer porté à la température de Curie, au spectre $S(\lambda)$ des neutrons incidents. Le rapport

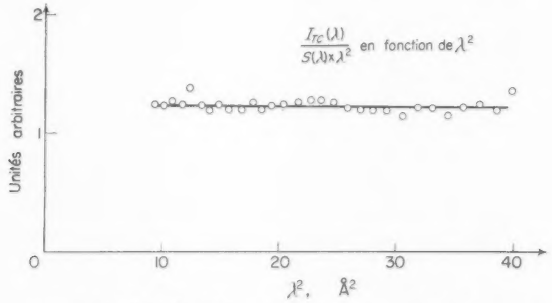


FIG. 2. $I_{T_C}(\lambda)/S(\lambda)\times\lambda^2$ en fonction du carré de la longueur d'onde.

$I_{T_C}(\lambda)/S(\lambda)\times\lambda^2$ est indépendant de la longueur d'onde (Fig. 2).

On peut fixer une limite supérieure à $(\kappa_1)_{T_C}$; on trouve que les corrélations à la température critique ont une portée supérieure à 70 Å.

D'autre part, la détermination de la valeur absolue de la section efficace $d\sigma/d\Omega$ à un angle de diffusion donné et pour une longueur d'onde donnée permet de déduire la valeur de $(\mu_1)_{T_C}$, à la température critique. Nous avons trouvé:

$$(\mu_1)_{T_C} = 1{,}46\pm0{,}17 \text{ Å}$$

(b) *Études à des températures supérieures à la température de Curie*

On mesure de la même façon le spectre $I_T(\lambda)$ des neutrons diffusés par le fer chauffé à une température $T>T_C$. On le compare alors au spectre diffusé dans les mêmes conditions et pendant le même temps à la température de Curie. Il est préférable d'utiliser des variations relatives avec la température; on évite de cette façon les erreurs sur la mesure de la valeur absolue de la section efficace. Comme nous avons précédemment vérifié la forme de la section efficace à la température de Curie, ceci n'introduit pas d'hypothèse.

D'après la formule de VAN HOVE, le rapport $I_{T_C}(\lambda)/I_T(\lambda)$ est égal à:

$$\frac{I_{T_C}(\lambda)}{I_T(\lambda)} = \frac{(\mu_1^2)_T}{(\mu_1^2)_{T_C}}\left[1+\frac{\kappa_1^2}{K^2}\right] = \frac{(\mu_1^2)_T}{(\mu_1^2)_{T_C}}\left[1+\frac{\kappa_1^2\lambda^2}{4\pi^2\theta^2}\right]$$

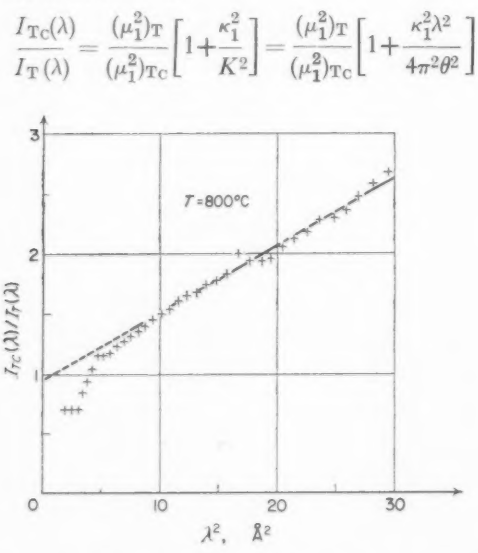


FIG. 3. Variation en fonction du carré de la longueur d'onde du rapport $I_{T_C}(\lambda)/I_T(\lambda)$ pour $T = 800°C$.

Nous avons vérifié que cette formule représente bien la variation de ce rapport (Fig. 3). Pour des longueurs d'ondes supérieures à 3 Å, le rapport $I_{T_C}(\lambda)/I_T(\lambda)$ a une variation linéaire en fonction du carré de la longueur d'onde.

L'écart à la loi linéaire aux courtes longueurs d'onde est attribué au comportement de la fonction de la corrélation au voisinage de l'origine; la formule de VAN HOVE, en effet, est obtenue en remplaçant la fonction de corrélation par son expression asymptotique, expression qui n'est pas valable au voisinage de l'origine. La section efficace s'écarte donc de la forme de VAN HOVE pour les grandes valeurs de K, c'est-à-dire les faibles longueurs d'onde.

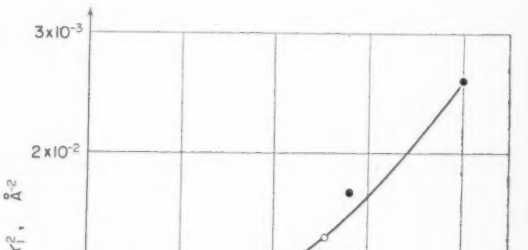


FIG. 4. Valeurs de κ_1^2
○ valeurs déterminées avec un spectre large (Chap. 2).
● valeurs déterminées avec un spectre étroit (Chap. 3).

La droite représentant la variation du rapport $I_{T_C}(\lambda)/I_T(\lambda)$ permet d'obtenir:

— la variation de μ_1 avec la température d'après son ordonnée à l'origine,

— la valeur de κ_1 à la température T, d'après sa pente.

Nous avons étudié des températures allant jusqu'à 50°C au-dessus de la température critique.

Les résultats sont donnés dans les Figs. 4 et 5 où sont portées, en fonction de la température, des valeurs de κ_1^2 d'une part et du rapport $(\mu_1^2)_T/(\mu_2^1)_{T_C}$ d'autre part.

Dans le tableau suivant, nous donnons les

VOL. 13 1960

valeurs de $1/\kappa_1$ c'est-à-dire la portée des corrélations aux diverses températures étudiées. L'erreur sur la portée est estimée à 15 pour cent; une partie provient de l'incertitude sur l'angle de diffusion, le reste provient des erreurs statistiques.

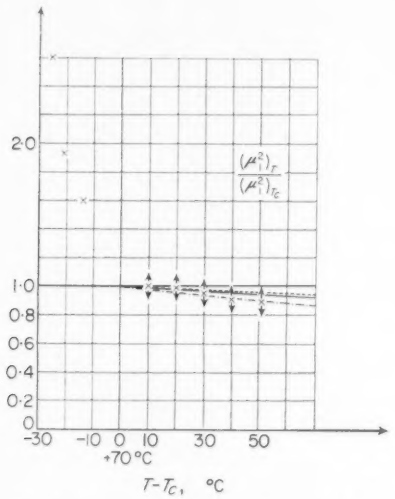


FIG. 5. Variation en fonction de la température du rapport $\mu_{1T}^2/\mu_{1T_C}^2$

× points expérimentaux
courbes calculées
— — — ELLIOTT (Bethe–Peierls 1ers voisins)
— · — · — DE GENNES (Normalisation à l'origine 1ers voisins)
——— DE GENNES (Normalisation à l'origine 2èmes voisins)

Tableau 1

Température (°C)	$1/K_1$ (Å)
780	25
790	16,7
800	12,5
810	10,2
821	8,8

Ces résultats sont en désaccord avec ceux de WILKINSON[5,6] qui trouve des portées environ une fois et demie plus faibles (Fig. 6). Sa méthode est différente de la nôtre. Il étudie la distribution angulaire des neutrons diffusés et en déduit la fonction de corrélation, par transformation de Fourier. Cette méthode, plus satisfaisante en principe, nécessite la connaissance complète de la distribution angulaire en particulier dans des régions de faible intensité. Ceci amène à utiliser un détecteur de grande dimension; les corrections

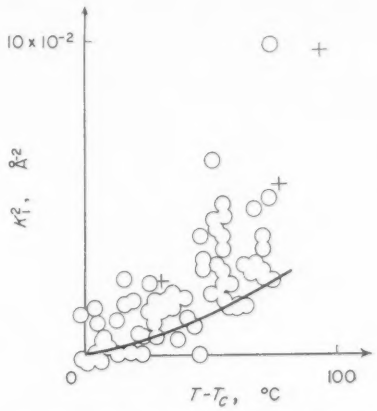


FIG. 6. Comparaison des valeurs de κ_1^2 obtenues par les différents expérimentateurs

\+ WILKINSON
○ LOWDE
— Valeurs obtenues à Saclay

qui en résultent deviennent importantes. Il est possible que ce soit l'origine du désaccord. LOWDE[7] obtient des valeurs intermédiaires.

(c) Comparaison des résultats avec les mesures magnétiques

On compare les valeurs de μ_1 obtenues directement et d'après la relation $(\kappa_1\mu_1)^2 = \chi_0/\chi$.

Les mesures de la susceptibilité du fer au-dessus de la température de Curie ont été faites par NEEL,[8] POTTER[9] et SUCKSMITH[10]. Les mesures sont en assez bon accord mais NEEL trouve que la susceptibilité d'un même échantillon n'a pas une valeur reproductible dans le domaine de température qu'il étudie, entre 770°C et 790°C. Il y a donc une incertitude sur la susceptibilité. Il est cependant intéressant de comparer les valeurs de μ_1 obtenues par la mesure directe et par la relation entre κ_1 et μ_1.

1° Valeur de μ_1 à la température critique.

La mesure directe donne $(\mu_1)_{T_C} = 1{,}46$ Å.

L'extrapolation à la température de Curie des

valeurs de μ_1 obtenues d'après la relation $(\kappa_1\mu_1)^2 = \chi_0/\chi$ donne $(\mu_1)_{T_C} = 1{,}33$ Å, en bon accord avec la mesure précédente.

2° Variation de μ_1 avec la température.

Dans le tableau suivant, nous avons porté les valeurs du rapport $(\mu_1^2)_T/(\mu_1^2)_{T_C}$ obtenues par les deux méthodes, qui donnent des résultats en bon accord.

Tableau 2. *Valeurs de* $(\mu_1^2)_T/(\mu_1^2)_{T_C}$

Température (°C)	Mesure directe	Relation $(\kappa_1\mu_1)^2 = \chi_0/\chi$
780	1	0,92
790	0,99	0,97
800	0,96	0,94
810	0,91	0,92
821	0,88	0,89

Comparaison aux valeurs calculées:

(1) Valeur de μ_1 à la température de Curie.

$(\mu_1)_{T_C}$ a été calculé par ELLIOTT[11] et DE GENNES[12] dans diverses approximations. Les résultats sont donnés dans le tableau suivant. Ils diffèrent un peu suivant que l'on considère comme voisins magnétiquement actifs, les premiers ou les seconds voisins d'un atome de fer. Le fer α a une structure cubique centrée; les premiers voisins sont au nombre de 8, ils sont situés à 2,48 Å. Les seconds voisins, au nombre de 6, légèrement plus éloignés, sont situés à 2,86 Å. NEEL[13] a suggéré que seuls les seconds voisins contribuent au magnétisme du fer, les premiers auraient une intégrale d'échange nulle.

Tableau 3. *Valeurs de* $(\mu_1)_{T_C}$ *en* Å

	Premiers voisins	Seconds voisins
Méthode du champ moléculaire	1,01	1,17
Approximation de BETHE–PEIERLS	1,09	1,31
Normalisation à l'origine	1,18	1,40

La méthode du champ moléculaire et l'approximation de BETHE–PEIERLS donnent des valeurs de $(\mu_1)_{T_C}$ qui n'assurent pas la condition de normalisation à l'origine $\gamma(o) = S(S+1)$. DE GENNES et GRANDIDIER[14] ont calculé les valeurs de $(\mu_1)_{T_C}$ telles que $\gamma(R)$ satisfait à la condition de normalisation à l'origine. Ces valeurs sont données dans la troisième ligne du tableau.

La valeur expérimentale obtenue est 1,46 Å par la mesure directe, 1,33 Å par la relation entre κ_1 et μ_1. L'accord avec les valeurs calculées est bon si on suppose les seconds voisins magnétiquement actifs. Nos résultats seraient donc en faveur de l'hypothèse de NEEL. Malheureusement les calculs ne sont pas assez sûrs pour permettre une conclusion. Les méthodes utilisées ne sont pas sûres tant du point de vue du traitement statistique que du point de vue du modèle physique qui suppose une interaction différente de zéro entre un seul type de voisins. Ces défauts semblent être effectifs car les méthodes utilisées ne rendent pas compte de la susceptibilité expérimentale du fer au voisinage supérieur de la température critique. La méthode du champ moléculaire donne une valeur de la susceptibilité environ deux fois et demie trop faible; avec l'approximation de BETHE–PEIRLS l'écart est réduit à un facteur deux mais il demeure encore très important.

(2) Variation de μ_1 avec la température.

Dans la Fig. 5 nous avons porté en même temps que les valeurs expérimentales les courbes calculées dans les différents modèles; toutes ces méthodes donnent des variations sensiblement égales, en bon accord avec celle trouvée expérimentalement. La variation de μ_1 avec la température n'est pas un bon test pour décider du meilleur modèle. Il faudrait améliorer beaucoup la précision expérimentale pour distinguer entre les différentes variations.

(d) Températures inférieures à T_C

Nous avons effectué la même étude pour des températures inférieures à la température critique. Nous trouvons que μ_1 varie beaucoup plus rapidement au-dessous de la température de Curie qu'au-dessus (Fig. 5). En réalité, il faut tenir compte des fluctuations d'aimantation longitudinales, c'est-à-dire dirigées suivant l'aimantation spontanée, et transversales. Nous étudions seulement une valeur moyenne μ_{1m} telle que

$$\frac{1}{\mu_{1m}^2} = \frac{2}{3}\frac{1}{\mu_{1l}^2} + \frac{1}{3}\frac{1}{\mu_{1t}^2}$$

VOL. 13 1960

l'indice t se rapporte aux fluctuations transversales, l'indice l aux fluctuations longitudinales.

3. ÉTUDE DE LA CINETIQUE DES FLUCTUATIONS

VAN HOVE a suggéré que les fluctuations d'aimantation évoluent dans le temps suivant une équation de diffusion

$$\frac{\partial \boldsymbol{M}}{\partial t} = \Lambda \, \nabla^2 \boldsymbol{M}.$$

La section efficace différentielle par unité d'angle solide et unité d'énergie des neutrons diffusés est alors donnée par l'expression:

$$\frac{d^2\sigma}{d\Omega d\epsilon} = \left(\frac{ge^2}{m_0c^2}\right)^2 \frac{2}{3\pi\hbar} \frac{S(S+1)}{\mu_1^2} \frac{1}{K^2+\kappa_1^2} \frac{\Lambda K^2}{\omega^2+\Lambda^2K^4}$$

A la température de Curie, il y a équilibre entre l'influence du champ moléculaire qui tend à augmenter l'alignement et l'influence de l'agitation thermique qui tend à le rompre; le temps de relaxation est infini, Λ est nul et la diffusion élastique. A des températures voisines les changements moyens d'énergie sont très petits; on ne peut obtenir un spectre de neutrons suffisamment étroit en énergie pour être capable de mesurer la forme exacte en fonction de l'énergie, de la section efficace $d^2\sigma/d\Omega d\epsilon$. On se contente d'utiliser un spectre suffisamment étroit pour que sa forme soit modifiée par l'inélasticité de la diffusion. La mesure de Λ est ainsi indirecte. Le principe des mesures est le même que celui des expériences précédentes: on envoie un faisceau de neutrons sur l'échantillon de fer chauffé, on étudie le spectre des neutrons diffusés à un angle de quelques degrés en fonction de la température de l'échantillon.

Les neutrons incidents sont rendus approximativement monocinétiques par un sélecteur mécanique[15], qui hache en même temps le faisceau, ce qui permet de mesurer le temps de vol des neutrons.

Le spectre en énergie des neutrons diffusés à une température T dépend du paramètre d'espace κ_1 et du paramètre de temps Λ (on suppose μ_1^2 inversement proportionnel à la température absolue).

On cherche par une méthode de moindres carrés les valeurs de ces paramètres. En fait, il est inutile d'introduire κ_1 comme un paramètre inconnu car sa valeur peut être déterminée *a priori*. En effet, l'aire totale du spectre diffusé ne dépend que de κ_1 et non de Λ. Ceci est rigoureusement vrai si les changements d'énergie sont assez petits pour que K soit indépendant de ω; l'intégrale de la section efficace sur les énergies finales $\int(d^2\sigma/d\Omega d\epsilon)\, d\epsilon$ est alors indépendante de Λ. Ce résultat reste une très bonne approximation si K varie un peu avec ω. Les valeurs de κ_1 déterminées ainsi sont en bon accord avec celles que nous avions précédemment mesurées (Fig. 4).

Il suffit de faire une méthode de moindres carrés pour déterminer Λ. Il faut d'abord calculer le spectre diffusé pour diverses valeurs de Λ. Ces calculs nécessitent l'emploi d'un calculateur électronique; chaque longueur d'onde du spectre diffusé reçoit une contribution de toutes les longueurs d'onde présentes dans le spectre incident. On fait une méthode de moindres carrés pondérés pour tenir compte des erreurs statistiques différentes sur les divers points expérimentaux.

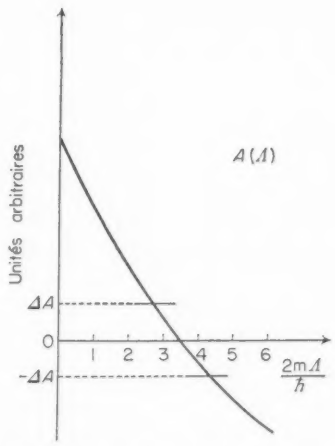


FIG. 7. Variation de l'aire $A(\Lambda)$ en fonction de Λ.

La Fig. 7 représente le spectre calculé dans l'hypothèse d'une diffusion élastique, ainsi que les points du spectre expérimental.

L'erreur sur Λ due aux erreurs statistiques sur les points expérimentaux est plus facile à évaluer si on détermine Λ par une méthode autre que la méthode des moindres carrés, mais qui donne pratiquement la même valeur de Λ. Les spectres calculés pour les différentes valeurs de Λ se croisent

tous en deux points. On cherche alors quelle est la valeur de Λ qui annule l'aire $A(\Lambda)$ comprise entre le spectre calculé et le spectre expérimental. Cette aire est composée de trois parties, une partie centrale et des ailes, dont on ajoute les valeurs absolues. La somme algébrique de ces trois quantités est nulle, puisque tous les spectres ont la même aire totale. On affecte à $A(\Lambda)$ un signe différent suivant que le spectre expérimental est dans la région centrale, au-dessous ou au-dessus du spectre calculé.

La Fig. 8 représente $A(\Lambda)$ en fonction de Λ. Il est alors facile de calculer l'erreur sur Λ, on

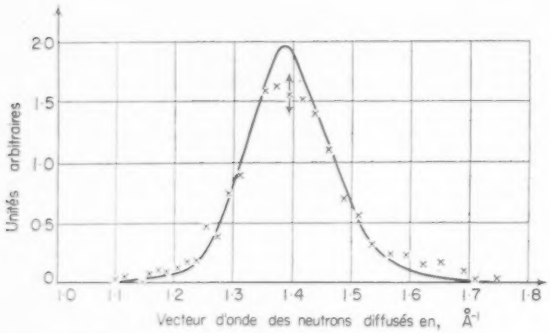


FIG. 8. Spectre des neutrons diffusés à la température de 850°C

× points expérimentaux
— courbe calculée en supposant la diffusion élastique

connait l'erreur $\triangle A$ sur $A(\Lambda)$ qui est l'erreur statistique sur l'aire totale du spectre expérimental. On en déduit l'erreur statistique sur Λ.

Les spectres calculés l'ont été à partir du spectre diffusé à la température de Curie, et non à partir du spectre des neutrons incidents. Cette méthode a l'avantage d'isoler les mécanismes de relaxation qui conservent le spin total.

En effet, dans les processus de relaxation envisagés par VAN HOVE, on ne tient compte que de ceux qui conservent le spin total; il existe également la relaxation spin réseau qui ne conserve pas le spin total. Au terme de relaxation ΛK^2 s'ajoute alors un terme $1/\tau$. La diffusion à la température critique n'est plus rigoureusement élastique, il est difficile d'avoir une estimation *a priori* du temps τ. Mais on peut montrer qu'avec notre méthode nous mesurons en fait la différence entre les constantes des temps μ la température T et à la température critique:

$$\left[\Lambda K^2+\left(\frac{1}{\tau}\right)_T\right]-\left(\frac{1}{\tau}\right)_{T_C}$$

On peut admettre que le temps τ ne dépend pas de la température puisque la relaxation spin réseau ne présente pas de singularité à la température critique et qu'on se limite à un intervalle étroit autour de T_c

$$\frac{T-T_c}{T_c}<\frac{1}{10}$$

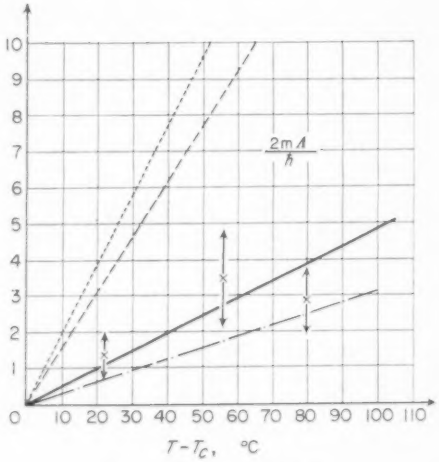


FIG. 9. Valeurs de $2m\Lambda/\hbar$

× points expérimentaux
—·— courbe calculée. Modèle de Heisenberg, 1ers voisins
———— courbe calculée. Modèle de Heisenberg, 2èmes voisins
— — — courbe calculée. Modèle des bands, 1ers voisins
- - - - - courbe calculée. Modèle des bandes 2èmes voisins

Sur la Fig. 9 sont portées les valeurs expérimentales du paramètre sans dimensions $2m\Lambda/\hbar$, ainsi que les courbes calculées. Les calculs ont été faits par DE GENNES[12] dans différents modèles: le modèle de Heisenberg et un modèle simple de bandes dans lequel on ne tient pas compte des corrélations entre les positions des électrons porteurs de magnétisme. Nos résultats sont en bon accord avec les valeurs calculées dans le modèle de Heisenberg, ils excluent le modèle de bandes tout

au moins dans cette forme simple. MOTT[15] a suggéré que les électrons magnétiques du fer, qui a une structure cubique centrée, ne font pas partie d'une bande de conduction, contrairement à ce qui se passerait pour des métaux de transition à structure cubique à faces centrées. Il est normal dans ce cas que nos résultats soient en accord avec les valeurs calculées dans le modèle de Heisenberg.

Il sera intéressant d'étudier le temps de relaxation des fluctuations dans le cas du nickel où l'on pense que les électrons magnétiques font partie d'une bande de conduction et les temps de relaxation peuvent être dans ce cas sensiblement différents. La section efficace de diffusion critique du nickel est d'après SQUIRES environ 50 fois plus petite que celle du fer et son étude nécessite de hauts flux de neutrons.

VOL. 13 1960

BIBLIOGRAPHIE

1. PALEVSKY H. et HUGHES D. J., *Phys. Rev.* **92**, 202 (1956).
2. SQUIRES G. L., *Proc. Phys. Soc., Lond.* **67**, 248 (1954).
3. VAN HOVE L., *Phys. Rev.* **95**, 249 (1954).
4. VAN HOVE L., *Phys. Rev.* **95**, 1374 (1954).
5. WILKINSON M. K. et SHULL C. G., *Phys. Rev.* **103**, 516 (1956).
6. GERSH H. A., SHULL C. G. et WILKINSON M. K., *Phys. Rev.* **103**, 525 (1956).
7. LOWDE R. D., *Rev. Mod. Phys.* **30**, 69 (1958).
8. NEEL L., *Ann. Phys.* **18**, 5 (1932).
9. POTTER H. H., *Proc. Roy. Soc.* A**146**, 362 (1934).
10. SUCKSMITH W. et PEARCE R. R., *Proc. Roy. Soc.* A**167**, 189 (1938).
11. ELLIOTT R. J. et MARSHALL W., *Rev. Mod. Phys.* **30**, 75 (1958).
12. DE GENNES P. G., Thèse, Université de Paris (1957); DE GENNES P. G. et HERPIN A., *C.R. Acad. Sci., Paris* **243**, 1611 (1956).
13. NEEL L., Congrès de magnétisme, Strasbourg, (1939).
14. GRANDIDIER C. *Rapport S.P.M.* nº 456, *Centre Études Nucléaires de Saclay.*
15. GOBERT G. et JACROT B., *J. Phys. Radium* **19**, 51A (1958).
16. MOTT N. F. et STEVENS K. W. H., *Phil. Mag.* 2, 1369 (1957).

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 13. pp. 244–250. Printed in Great Britain.

A METHOD OF EVALUATING DIFFUSION COEFFICIENTS IN CRYSTALS*

O. P. MANLEY

Department of Electrical Engineering and Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts

(*Received* 24 *August* 1959; *revised* 23 *October* 1959)

Abstract—The mechanism of diffusion proposed by RICE is reconsidered with the aid of an extension of KAC's theorem. The analysis yields an activation energy which, as predicted by ZENER, is simply related to the minimal local deformation energy. It is also found that in contrast with RICE's treatment, the activation energy may be calculated directly from the atomic force constants without resorting to normal mode analysis.

INTRODUCTION

RECENTLY RICE[1] suggested a method of evaluating the parameters entering into the diffusion coefficient of particles in a crystal lattice. This method rests on the following mechanism for the migration of the diffusing particle: It is assumed that the particle moves irreversibly to an appropriate neighboring site whenever simultaneously (i) the said site is vacant, (ii) the particle has achieved a critical displacement along a line joining the old site with the new site, and (iii) all the other particles in the lattice which might conceivably impede the migration happen to be out of the way. It is further assumed that the harmonic approximation is valid and that quantum effects may be neglected.

The mechanism summarized above is sufficiently general to apply to all the conceivable models for diffusion. To fix ideas the case of interstitial diffusion will be referred to below. In this case then, whenever the concentration of the diffusing species is very small, condition (i) is practically always certain to be fulfilled, and hence the effects of the contrary situation will be neglected hereafter.

Following RICE, the frequency with which conditions (ii) and (iii) are met simultaneously is denoted by the symbol $P(\{\delta\})$. RICE evaluates $P(\{\delta\})$ subject to the following simplifying assumptions[1, 2]: first, $P(\{\delta\})$ is represented as the product of the frequency with which (ii) occurs and the probability of (iii) occurring simultaneously; and secondly, it is assumed that it suffices to compute the probability of the occurrence of the minimal configuration satisfying condition (iii). Specifically, RICE assumes that since the minimal favorable configuration—i.e. the configuration in which (ii) and (iii) are just barely fulfilled—is very improbable, the other favorable configurations are even less probable and hence their contributions to the diffusion coefficient are negligible.

One of the drawbacks of RICE's approach is that there is no way within the framework of the method by which the effect of the two simplifications can be estimated; hence it appears impossible to improve the approximation within the domain of the method.

An alternate approach which does not seem to suffer from the above mentioned disadvantage and which is subject to less drastic simplifications is presented below. The method consists of a direct calculation of the frequency with which conditions (ii) and (iii) are fulfilled. The calculation is based on a simple extension of KAC's theorem[3] and consists of two parts. First, we compute the

* This paper is based in parts on a thesis to be submitted in partial fulfilment of the requirements for the degree of Doctor of Science or Doctor of Philosophy in the Department of Electrical Engineering at the Massachusetts Institute of Technology on or about May 1960.

asymptotic number of jumps per unit time for a fixed distribution of energies of the normal modes of the crystal. Secondly, we compute the approximate thermal average of that number and equate it to the frequency $P(\{\delta\})$. It is then shown that in the lowest approximation, the contribution to the diffusion activation energy arising from $P(\{\delta\})$ is simply related to the elastic energy associated with the local lattice deformation during the migration of the diffusing particle.

CALCULATION OF THE FREQUENCY FACTOR $P\{\delta\}$

VOL. 13 1960

It is assumed that a normal mode analysis has been carried out and that the effects of the interstitial particles have been included in the analysis. Then it follows that in a crystal bound with harmonic forces, for a given distribution of normal mode energies, the displacement from equilibrium, F_j, of the j^{th} atom is given by

$$F_j = \sum_{k=1}^{N} \alpha_{jk} \epsilon_k^{1/2} \cos(\omega_k t + \delta_k) \tag{1}$$

where $\sum_k \epsilon_k = U$ is the internal energy of the crystal, α_{jk}'s are determined by the normal mode analysis, N is the number of degrees of freedom of the system, ω_k is the frequency of the k^{th} mode and δ_k is a phase angle.

Consider an N-dimensional space with orthogonal coordinates θ_k such that

$$0 \leqslant \theta_k < 1 \quad k = 1,2 \ldots N \tag{2}$$

Then we can set $(\omega_k t + \delta_k) = 2\pi\theta_k \pmod{2\pi}$. Thus, for a given set of normal mode energies, ϵ_k, the displacement from equilibrium, F_j, may be completely described by the set of independent vectors, θ_k, rather than by the single variable t. This procedure may be shown to be valid[5] whenever the frequencies ω_k form a linearly independent set; which is usually true in the case of normal modes of a crystal. We now want to find the average fraction of time during which simultaneously the migrating particle achieves a critical displacement from equilibrium q_{10}, i.e.

$$q_{10} - \frac{dq_{10}}{2} < F_1 < q_{10} + \frac{dq_{10}}{2}$$

with some velocity

$$\dot{q}_{10} - \tfrac{1}{2}\, d\dot{q}_1 < \frac{dF_1}{dt} < q_1 + \tfrac{1}{2}\, d\dot{q}_1$$

and all the other particles have displacements F_j in the intervals $(q_{j0},\ q_{j0}+dq_{j0})$. We denote this fraction of time by

$$M(q_{10},\, q_{20} \ldots q_{n0},\, \dot{q}_1)\, d\dot{q}_1 \prod_{j=1}^{n} dq_{j0}$$

In the N-dimensional space the required fraction of time M is represented asymptotically by a volume element which may be found by integrating over the variables θ_k;

$$M(q_{10} \ldots q_{n0},\, \dot{q}_1)\, d\dot{q}_1 \prod_{j=1}^{n} dq_{j0}$$

$$= d\dot{q}_1 \prod_{j=1}^{n} dq_{j0} \int \ldots \int d\theta_1 \ldots d\theta_N \tag{3}$$

The limits of integration are determined by the conditions on F_j's specified above. It can be shown[3, 4] that equation (3) reduces to

$$M(q_{10},\, q_{20},\, \ldots q_{n0},\, \dot{q}_1)\, d\dot{q}_1 \prod_{j=1}^{n} dq_{j0}$$

$$= d\dot{q}_1 \prod_{j=1}^{n} dq_{j0} \left(\frac{1}{2\pi}\right)^n \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} d\lambda \prod_{j=1}^{n} d\eta_j \times$$

$$\times \int_0^1 \ldots \int_0^1 \prod_{k=1}^{N} d\theta_k \times$$

$$\times \exp\left\{ i \left[\sum_{j=1}^{n} \eta_j (F_j - q_{j0}) + \lambda (F_1' - \dot{q}_1) \right] \right\} \tag{4}$$

where q_{10} is the critical displacement of the migrating atom, $\dot{q}_1$ is its velocity $F_1' = dF_{1/dt}$ and q_{j0}, $j \neq 1$, are the displacements of the other particles. In order to determine the number of times per

unit time that the conditions (b) and (c) are satisfied, and hence the number of times per unit time that a jump occurs, consider the space spanned by the n vectors q_{j0}. The actual time spent in the neighborhood of the point specified above in this space is given by

$$\frac{dq_{10}}{|\dot{q}_1|} \tag{5}$$

Therefore, the asymptotic number of times per unit time that the specified configuration occurs is the ratio of equations (4) to (5) integrated over the velocity $\dot{q}_1$. Now the, configuration favorable to the vacancy jump occurs over a range of values q_{j0}, $j \geqslant 2$, say $q_{j\min} < q_{j0} < q_{j\max}$. Thus L, the total number of jumps per unit time, is given by:

$$L = \frac{1}{2}\left(\frac{1}{2\pi}\right)^n \int_{q_{2\min}}^{q_{2\max}} \cdots \int_{q_{n\min}}^{q_{n\max}} \prod_{j=2}^{n} dq_{j0} \int_{-\infty}^{\infty} d\dot{q}_1 \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} d\lambda \prod_{j=1}^{n} d\eta_j \int_0^1 \cdots \int_0^1 \prod_{k=1}^{N} d\theta_k \times$$

$$\times \exp\left\{ i \left[\sum_{j=1}^{n} \eta_j (F_j - q_{j0}) + \lambda(F_1' - \dot{q}_1) \right] \right\} |\dot{q}_1| \tag{6}$$

The factor of $\frac{1}{2}$ occurs since only the approach to q_{10} from below is of interest.

It is now noted that

$$|F_1'| = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{1-\cos F_1'\xi}{\xi^2} d\xi \tag{7}$$

Therefore, after changing the order of integration and integrating over λ, $\dot{q}$ and q_{j0}'s, we have

$$L = \left(\frac{1}{2\pi}\right)^{n+1} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \prod_{j=1}^{n} d\eta_j \frac{d\xi}{\xi^2} \int_0^1 \cdots \int_0^1 \prod_{k=1}^{N} d\theta_k \times$$

$$\times \exp(-i q_{10}\eta_1) \exp\left[i \sum_{j=1}^{n} \eta_j F_j (1-\cos F_1'\xi) \right] \times$$

$$\times \prod_{j=2}^{n} \frac{\exp(-i\eta_j q_{j\max}) - \exp(-i\eta_j q_{j\min})}{-i\eta_j} \tag{8}$$

The maximum possible displacement of the j^{th} particle for a given set of normal modes is

$$q_{j\max} = \sum_k |\alpha_{jk}| \epsilon_k^{1/2} \tag{9}$$

We recognize two limiting values of $q_{j\min}$. First, that corresponding to the situation where q_{j0} must be close to $q_{j\max}$ in order for a favorable configuration to exist. This is the case for some of the nearest neighbors of the migrating atom; namely, those which, when in their equilibrium position, would impede the migration: in this case we can assume

$$q_{j\min} = q_{j\max}\left(1 - \frac{\Delta q_j}{q_{j\max}}\right) \tag{10}$$

Second is the limiting value corresponding to the situation where the atom can be in nearly every portion of its swing for a favorable configuration to exist. Here we can assume

$$q_{j\min} = -q_{j\max}\left(1 - \frac{\Delta q_j}{q_{j\max}}\right) \tag{11}$$

Assume that displacements q_{j0}, $j = 2,3, \ldots m$ belong to the first class while those for $j > m$ belong to the second class. For the sake of clarity we shall restrict our attention to the first class. In other words, we assume that the rate is limited by these displacements alone. The extension of the arguments to the second class will become obvious.

$$L = \left(\frac{1}{2\pi}\right)^{m+1} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \prod_{j=1}^{m} d\eta_j \frac{d\xi}{\xi^2} \int_0^1 \cdots \int_0^1 \prod_{k=1}^{N} d\theta_k \exp[i\eta_1(F_1 - q_{10})] \cos(1 - F_1'\xi) \times$$

$$\times \exp\left[i \sum_{j=2}^{m} \eta_j (F_j - q_{j\min}) \right] \prod_{j=2}^{m} \Delta q_j \sum_{t=0}^{\infty} \frac{(-i\eta_j \Delta q_j)^t}{(t+1)!}$$

$$= \left(\prod_{j=2}^{m} \Delta q_j \right) \sum_{t_2, t_3, \cdots t_m = 0}^{\infty} \prod_{r=2}^{m} \frac{1}{(t_r+1)!} \left(\Delta q_r \frac{\partial}{\partial q_{r\min}} \right)^{t_r} I(q_{10}, q_{2\min}, \cdots q_{m\min}) \tag{12}$$

where

$$I(q_{10}, q_{2\min}, \cdots q_{m\min}) = \left(\frac{1}{2\pi}\right)^{m+1} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \prod_{j=1}^{m} d\eta_j \frac{d\xi}{\xi^2} \exp(-i\eta_1 q_{10}) \exp\left(-i \sum_{j=2}^{m} \eta_j q_{j\min} \right) \times$$

$$\times \left(\prod_{k=1}^{N} J_0 \left[\epsilon_k^{1/2} \sum_{j=1}^{m} \eta_j \alpha_{jk} \right] - \prod_{k=1}^{N} J_0 \left\{ \epsilon_k^{1/2} \sqrt{\left[\left(\sum_{j=1}^{m} \eta_j \alpha_{jk} \right)^2 + \xi^2 \omega_k^2 \alpha_{1k}^2 \right]} \right\} \right) \tag{13}$$

So far we have calculated the number of jumps per unit time for a unique distribution of normal mode energies. To obtain the frequency $P(\{\delta\})$, it is necessary to evaluate the average of L over a canonical ensemble. Thus

$$P(\{\delta\}) = \int_0^{\infty} \cdots \int_0^{\infty} L \times$$

$$\times \exp\left(-\frac{1}{kT} \sum_{k=1}^{N} \epsilon_k \right) \prod_{k=1}^{N} d\left(\frac{\epsilon_k}{kT}\right) \tag{14}$$

where k is the Boltzmann constant and T is the absolute temperature.

It is found that an exact closed expression for this ensemble average cannot be obtained. However, if it is assumed that Δq_j may be replaced by its most probable value, which in turn is set equal to the ensemble average of $\Delta q_j \equiv \overline{\Delta q_j}$, $j = 2, 3, \ldots m$, then a remarkably simple expression for $P(\{\delta\})$ results.

To simplify notation, we shall use q_j with the understanding that $q_1 = q_{10}$, $q_j = q_{j\min}$, $j \neq 1$.

Note that $q_{j\min}$ is not a function of energy, but depends merely on the geometry of the lattice and the relative sizes of the migrating atom and the lattice parameters. Then the thermal average of $I(q_j)$ is given by

$$\overline{I(q_j)} = \int_0^{\infty} \cdots \int_0^{\infty} I(q_j) \times$$

$$\times \exp\left(-\frac{1}{kT} \sum_{k=1}^{N} \epsilon_k \right) \prod_{k=1}^{N} d\left(\frac{\epsilon_k}{kT}\right) =$$

$$= \left(\frac{1}{2\pi}\right)^{m+1} \sqrt{\left(\pi kT \sum_{k=1}^{N} \alpha_{1k}^2 \omega_k^2 \right)} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \prod_{j=1}^{m} d\eta_j \times$$

$$\times \exp\left[-i \sum_{j=1}^{m} \eta_j q_j - \right.$$

$$\left. - \frac{kT}{4} \sum_{k=1}^{N} \left(\sum_{j=1}^{m} \eta_j \alpha_{jk} \right)^2 \right] \tag{15}$$

To carry out the indicated integrations we set

$$A_{ij} = \sum_{k=1}^{N} \alpha_{ik}\,\alpha_{jk} \tag{16}$$

And let

$$\sum_{i,j}^{m} A_{ij}\,\eta_i\,\eta_j = (a_{11}\,\eta_1 + a_{12}\,\eta_2 \ldots a_{1m}\,\eta_m)^2 +$$

$$+ (a_{22}\eta_2 \ldots a_{2m}\eta_m)^2 + a_{mm}^2\eta_m^2 = \sum_{j=1}^{m} \psi_j{}^2 \tag{17}$$

where

$$\psi_j = \sum_{i=j}^{m} a_{ji}\eta_i \tag{18}$$

and

$$\eta_j = \sum_{i=j}^{m} (a^{-1})_{ji}\,\psi_i = \sum_{i=j}^{m} b_{ji}\,\psi_i \tag{19}$$

Note that $[A] = [a]^T[a]$ and therefore the Jacobian

$$J\left[\frac{\partial(\eta_j)}{\partial(\psi_j)}\right] = \frac{1}{\sqrt{(\det\,[A])}} \tag{20}$$

Thus the change of variables, equation (19), leads to

$$\overline{I(q_j)} = \left(\frac{1}{2\pi}\right)^{m+1} \frac{\sqrt{\left(\pi kT \sum_{k=1}^{N} \alpha_{1k}^2 \omega_k^2\right)}}{\sqrt{(\det\,[A])}} \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} \prod_{j=1}^{m} d\psi_j \exp\left[-i \sum_{j}^{m} \psi_j \sum_{i=1}^{j} q_i b_{ij}\right] \times$$

$$\times \exp\left[-\frac{kT}{4} \sum_{j=1}^{m} \psi_j^2\right]$$

$$= \frac{\sqrt{\left(\sum_{k=1}^{N} \alpha_{1k}^2\,\omega_k^2\right)}}{2\pi\sqrt{(\det\,[A])}} \left(\frac{1}{\sqrt{(\pi kT)}}\right)^{m-1} \exp\left[-\frac{1}{kT} \sum_{i,k}^{m} (A^{-1})_{ik} q_i q_k\right] \tag{21}$$

Hence, using equations (12) and (21) in equation (14), the jump frequency is found to be:

$$P(\{\delta\}) = \gamma_1 \frac{\sqrt{\left(\sum_{k=1}^{N} \alpha_{1k}^2\right)}}{\sqrt{(\det\,[A])}} \prod_{j=2}^{m} \frac{\overline{\Delta q_j}}{\sqrt{(\pi kT)}} \sum_{t_2, t_3, \ldots t_m = 0} \prod_{r=2}^{m} \frac{1}{(t_r+1)!} \left(\overline{\Delta q_r} \frac{\partial}{\partial q_r}\right)^{t_r} \times$$

$$\times \exp\left[-\frac{1}{kT} \sum_{i,k}^{m} (A^{-1})_{ik}\,q_i\,q_k\right]$$

$$= \gamma_1 \sqrt{\left(\frac{\sum_{k=1}^{N} \alpha_{1k}^2}{\det\,[A]}\right)} \int_{q_2}^{q_2 + \overline{\Delta q_2}} \frac{dq_2}{\sqrt{(\pi kT)}} \ldots \int_{q_m}^{q_m + \overline{\Delta q_m}} \frac{dq_m}{\sqrt{(\pi kT)}} \exp\left[-\frac{1}{kT} \sum_{i,k}^{m} (A^{-1})_{ik}\,q_i\,q_i\right]$$

$$= \gamma_1 \sqrt{\left(\frac{\sum_{k=1}^{N} \alpha_{1k}^2}{\det[A]}\right)} \prod_{j=2}^{m} \frac{\overline{\Delta q_j}}{\sqrt{(\pi kT)}} \exp\left[-\frac{1}{kT} \sum_{i,k}^{m} (A^{-1})_{ik}\,q_i\,q_k\right] + 0\left[\left(\frac{\overline{\Delta q_i}}{\sqrt{(\pi kT)}}\right)^m\right] \tag{22}$$

where

$$\gamma_1 = \frac{1}{2\pi}\sqrt{\left(\frac{\Sigma\alpha_{1k}^2\,\omega_k^2}{\Sigma\alpha_{1k}^2}\right)} \qquad (23)$$

It should be noted that if the inclusion of terms

$$0\left[\left(\frac{\overline{\Delta q_j}}{\sqrt{(\pi kT)}}\right)^m\right]$$

is contemplated, then terms dependent on the second class of displacements should be considered, since the latter are of the same order in

$$\frac{\overline{\Delta q_j}}{\sqrt{(\pi kT)}}$$

VOL. 13 1960

A few words about the second form of equation (22). It seems noteworthy to point out that the multiple integration has a simple physical meaning: since the configuration required for a jump is close to an extremum of motion, there is little kinetic energy associated with it. Hence, to a good approximation, the energy of interest is some potential energy only. The integrand is then the probability that a given configuration (requiring a specified potential energy) obtains at a given temperature, T. Since, by hypothesis, for a specified critical displacement of the migrating atom, the jump may occur over a range of configurations of the nearest neighbors, the multiple integration over that range gives the total probability of occurrence of a suitable configuration. It should be noted further that the frequency factor $P(\{\delta\})$ has the general form of a frequency of vibration of the migrating atom multiplied by the probability of occurence of a favorable configuration. This is consistent with the phenomenological arguments of ZENER and others.[6]

THE ACTIVATION ENERGY

To interpret the significance of the energy in the exponent of the exponential in equation (22), we note that for a harmonic crystal the total energy is given by

$$H = \frac{1}{2}\sum M_{rs}F_r'F_s' + \frac{1}{2}\sum K_{rs}F_rF_s \qquad (24)$$

where, as before, the primes indicate time differentiation, M_{rs} are the atomic mass constants, and K_{rs} are the atomic force constants. It should be emphasized that this Hamiltonian includes the effects of the migrating particles present in the crystal.

The normal mode analysis consistent with the representation of equation (1) leads to the following relationships:[5]

$$\tfrac{1}{2}\alpha^T\boldsymbol{M}\alpha = \wedge^{-1} \text{ and } \tfrac{1}{2}\alpha^T\overline{\boldsymbol{K}}\alpha = I \qquad (25)$$

where α^T is the transpose of the modal matrix α, $\wedge$ is the diagonal matrix of eigenvalues of the Hamiltonian and I is the identity matrix. Hence it follows that

$$A_{ij} = \sum_{k=1}^{m}\alpha_{ik}\alpha_{jk} = 2(\overline{K}^{-1})_{ij} \qquad (26)$$

Whence approximately

$$(A^{-1})_{ij} \approx \tfrac{1}{2}\,\overline{K}_{ij} \qquad (27)$$

Although the above approximation is not to be taken too seriously, it does show that the activation energy for particle diffusion is rather simply related to the elastic energy associated with the minimal local deformation necessary for the migration of the diffusing particle. In fact, it is noteworthy that to the lowest order in $\overline{\Delta q_r}$, the activation energy may be computed directly from the model of the crystal (inclusive of impurities) without resorting to normal mode analysis. Thus, given the matrix of the atomic force constants K_{rs}, the matrix is first inverted; then the set of elements of the inverted matrix with indices corresponding to the m degrees of freedom under consideration is partitioned from the remainder of the matrix; finally, the resulting $m\times m$ matrix is inverted, yielding the modified force constants $(A^{-1})_{ij}$. A set of q_j's consistent with the geometry of the lattice is chosen, and the deformation energy associated with these displacements is easily computed.

Physically this procedure is equivalent to assuming that all the atoms not directly connected with the migration have arbitrary displacements. Thus the whole crystal is replaced by the migrating atom coupled to some of its nearest neighbors with suitably modified coupling constants, this subsystem being immersed in a thermal bath made up of the rest of the crystal.

THE THERMODYNAMIC APPROXIMATION

We turn now to a brief discussion of the thermodynamic approximation involved in replacing Δq_j by $\overline{\Delta q_j}$. First we set

$$\Delta q_j = \delta_j q_{j\max}$$

$$= \delta_j \sum_{k=1}^{N} |\alpha_{jk}| \, \epsilon_k^{1/2}, \; j = 2, 3 \ldots m \qquad (28)$$

whence

$$\overline{\Delta q_j} = \delta_j \int_{U > U_0}^{\infty} \cdots \int^{\infty} \sum_{k=1}^{N} |\alpha_{jk}| \, \epsilon_k^{1/2} \times$$

$$\times \exp\left(-\sum_{i=l}^{N} \frac{\epsilon_i}{kT}\right) \prod_{i=1}^{N} d\left(\frac{\epsilon_i}{kT}\right) \qquad (29)$$

where

$$U = \sum_{i=1}^{N} \epsilon_i, \text{ and } U_0 = \frac{(q_{10})^2}{\sum_{k=1}^{N} \alpha_{1k}^2}$$

The restriction on the lower limits of integration arises from the requirement that $F_1 = q_{10}$, which in turn leads to the condition for the minimum energy of the normal modes.[1] However, it can be shown that apart from terms in the order of less than

$$\frac{1}{N!}\left(\frac{U_0}{kT}\right)^N$$

the lower limits can be extended to zero. Thus

$$\overline{\Delta q_j} = \frac{\delta_j}{2} \sqrt{(\pi k T)} \sum_{k=1}^{N} |\alpha_{jk}| \qquad (30)$$

In principle, it is now possible to estimate the error introduced by this approximation. For instance, one could find an appropriate expansion of equation (12), and then subtract from it the terms corresponding to equation (22). The remainder series could be bounded, yielding the required estimate. It is felt, however, that the results thus obtained may not merit the computational labor involved. To see that we argue as follows.

The thermal motion of the atoms in the lattice may be regarded as fluctuations of the position vectors of those atoms. Now, if appropriate thermal averages of the fluctuations are computed, and then one demands the deviations from those averages, then in effect one is asking for estimates of higher order fluctuations. Since in the present treatment, both quantum effects and anharmonicities have been neglected, and since these may become sufficiently important in higher order approximations, the significance of such approximations based on higher order fluctuations alone seems rather dubious.

CONCLUSION

In conclusion, it is clear that ZENER's hypothesis[6] concerning the nature of the activation energy for diffusion is supported to a large extent by the results presented above. On the other hand, it is equally clear that a significant contribution to the entropy of activation may arise from the pre-exponential factor. But even that part of the entropy depends primarily on the temperature dependence of the atomic force constants and hence on the elastic properties of the crystal.

Acknowledgement—The author wishes to acknowledge the stimulating discussions with Prof. R. B. ADLER of this laboratory, Prof. M. KAC of Cornell University and Prof. S. A. RICE of University of Chicago.

Part of this research was sponsored by ERDAFCRC, ARDC under Contract AF33(616)3984 and by ONR under Contract Nonr-1841(51). This paper, although based on work sponsored by the above mentioned agencies, has been neither approved nor disapproved by them. The work was also supported in part by Signal Corps Contract DA36-039-sc-78108.

REFERENCES

1. RICE S. A., *Phys. Rev.* **112**, 804 (1958).
2. MANLEY O. P. and RICE S. A., *Phys. Rev.*, to be published.
3. KAC M., *Amer. J. Math.* **65**, 609 (1943).
4. CHANDRASEKHAR S., *Rev. Mod. Phys.* **16**, 1 (1943).
5. SLATER N. B., *Proc. Roy. Soc.* A**194**, 112 (1948).
6. ZENER C., *Imperfections in Nearly Perfect Crystals* p. 289. Wiley, New York (1952).

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 13. pp. 251–256. Printed in Great Britain.

MECHANICALLY-EXCITED EMISSION IN CADMIUM SULFIDE

D. M. WARSCHAUER† and D. C. REYNOLDS

Aeronautical Research Laboratory, Wright-Patterson Air Force Base, Ohio

(*Received* 4 *August* 1959; *revised* 15 *October* 1959)

Abstract—The phenomenon of radiative emission obtained by mechanical excitation of optically or thermally stimulated cadmium sulfide is described and discussed in terms of the recently proposed band structure of the material and the existence of trapping levels in the forbidden-gap region.

VOL. 13 1960

INTRODUCTION

DESPITE the fact that the properties of cadmium sulfide have undergone intensive investigation for many years,[1] new effects are still being discovered and understanding of the material is becoming more complete as increasingly pure and perfect crystals are produced.

LAMBE *et al.*[2] presented evidence that the edge emission observed in cadmium sulfide results from recombination of a free hole with a trapped electron, in which the electronic transition involves longitudinal optical modes of the lattice. FURLONG[3] observed a weaker set of lines having the same spacing as the edge emission lines, but displaced 0·025 eV to the short-wavelength side. DUTTON[4] has examined the dichroism evident in the absorption edges of the material and the fact that the edge emission is polarized predominantly perpendicular to the *c*-axis of the crystal, independent of orientation. LEMPICKI[5] has verified this observation, and pointed out the failure of anisotropic oscillator models to account for this effect. KELLY and FREDERICKS[6] have measured the shift of the photoconductivity excitation peak as a function of the orientation of the plane of polarization of the exciting light and shown that like the absorption edges, but unlike fluorescent emission, the photoconductivity excitation is dependent on incident polarization. They point out the number of effects in CdS yielding an energy difference of the order of 0·02 eV.

These observations are accounted for by a model derived from that of LAMBE and KLICK[7] by BIRMAN[8] and HOPFIELD[9].

In this model the extrema of the valence and conduction bands are assumed to occur at $k = 0$, with the valence band consisting of three branches split by spin–orbit and crystal-field interactions. The absorption, luminescence and photoconductivity phenomena mentioned earlier can be interpreted in terms of transitions between the Γ_7 conduction band, the Γ_9 and Γ_7 valence bands, a trap level just below the conduction band and a set of recombination centers. The traps and recombination centers are assumed to have Γ_7 symmetry, or something close to it. Transitions of the form Γ_7–Γ_9 involve polarization $\perp$ *c*-axis while Γ_7–Γ_7 involve either $\parallel$ or $\perp$ polarization. The recombination center is assumed by BIRMAN to be 1·99 eV above the valence band, and thus to be responsible for 6200 Å fluorescence which should thus be resolvable into a polarized band in conjunction with an unpolarized one.

The effect on the emission of impurities such as silver and the existence of quenching bands were investigated earlier in some detail.[7, 10] Recently, WHEELER has used the symmetries of the model just discussed to explain the properties of exciton emission shorter than the wavelengths associated with the band gap.[11] On the other hand, COLLINS[12] prefers to invert the recombination mechanism of this model and the earlier LAMBE–KLICK–DEXTER model[2] so that edge emission is produced when a free electron recombines with a trapped hole.

† Now at Research Division, Raytheon Company, Waltham, Massachusetts.

It is our purpose here to introduce the experimental facts of mechanically-excited luminescence and to interpret these observations on the basis of the model described. This interpretation is not unique, and in fact one piece of evidence to be presented (emission due to Cu) is not at all in accord with the model generally accepted at present. Further, it should be pointed out that other evidence, namely that of preliminary experiments on cyclotron resonance[13] and magnetoresistance[14] indicate that the band structure may be considerably more complex than that indicated by the present model.

EXPERIMENTAL OBSERVATIONS

Not all cadmium sulfide crystals exhibit the effect to be described; in general, the purer and more perfect the crystal, the more likely it is to show mechanical excitation. (An exception to this is the effect observed in crystals which have been intentionally doped. This will be described later.) The observations to be discussed here were all made on large single crystals grown in this laboratory[15] and also independently elsewhere. The crystals are predominantly wurzite structure and, with the exception of intentionally-doped crystals, are as pure as possible.* Single crystals of cubic zinc sulfide have also been examined, but without success in detecting an effect. Single crystals of mixed zinc-cadmium sulfide, which are predominantly wurzite structure, do show the effect up to at least 15 per cent of zinc in cadmium sulfide. It is possible that the effect exists even beyond this composition, for example in hexagonal zinc sulfide, but has not been observed because of difficulties in detection.

Suppose a cadmium sulfide crystal is cooled, by immersion, to the temperature of liquid nitrogen. It is well known that stimulation of the crystal by irradiation with ultraviolet light causes emission of the green fluorescence known as "edge emission", which is usually identified with recombination of a trapped electron and a free hole.[2] If the crystal is raised to a temperature somewhat above nitrogen temperature, edge emission can no longer be stimulated. SIHVONEN *et al.*[16] have observed green emission at room temperature in some crystals. REYNOLDS[17] has shown that this emission corresponds to a temperature shift of the short-wavelength blue emission peaks observed at liquid nitrogen temperature. This is not considered to be edge emission here.

Suppose a cooled crystal which shows strong edge emission is first stimulated, say by irradiation with ultraviolet, and then completely shielded from light. If the crystal is now gently tapped, it is found that a green flash of light is emitted from the body of the crystal. Although the lower bound to the amount of mechanical excitation necessary to cause emission has not been investigated, it is known that impulses involving energies of 750 ergs or smaller are sufficient to excite the effect. The mechanical energy seems to have to be expended in the form of an impact; squeezing the crystal by unilateral compression slowly applied does not excite emission for pressures up to the breaking point of the crystal. Setting the crystal into ultrasonic vibration does not cause emission either. One reliable method of excitation is to drop spheres of known weight from a calibrated height onto the surface of the crystal.

It is found that the impact is most effective when applied parallel to the *c*-axis of the crystal, with excitation perpendicular to the axis having little or no effect. When examined in detail, the flash appears to consist of streamers of light emanating from the point of impact and radiating predominantly in a direction parallel to the *c*-axis. Spectrographic examination of the flashes is difficult, but photographic evidence indicates that the flashes are composed of most, if not all, of the spectral lines which comprise edge emission. In addition, it is found that the emission is polarized with the *E*-vector perpendicular to the *c*-axis, in accord with the direction of polarization observed for normal edge emission,[5] indicating that a Γ_7–Γ_9 transition appears to be operative.

Repeated tapping produces a flash for each tap. Of the order of a thousand flashes may be obtained from a well-stimulated crystal by tapping more or less uniformly over the whole surface. After the stimulation is exhausted, the crystal can be reactivated by brief illumination, and the process of obtaining multiple flashes by tapping can be repeated.

* GE luminescent grade CdS powder, from which these crystals were grown, was analyzed to have 5 p.p.m. Si, 1 p.p.m. Mg, 0·5 p.p.m. Cu, 50 p.p.m. Zn. The single crystal showed 5 p.p.m. Si, 0·5 p.p.m. Mg, with Cu, Zn, Ag non-detectable.

VOL.
13
1960

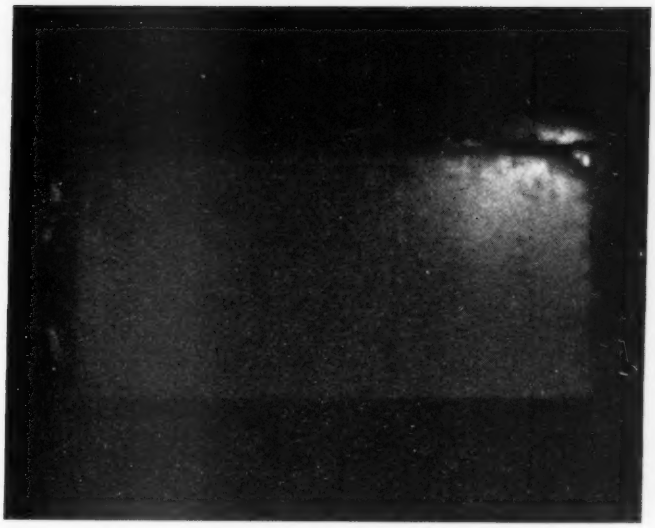


FIG. 1. Photograph of thermally-stimulated emission from rectangular parallelepiped of crystal tapped at nitrogen temperature. Crystal was tapped at brightest corner.

[*facing p.* 252

VOL
17
196

VOL. 13 1960

It is found that ultraviolet is not required for stimulation of the effect; in fact, the threshold for stimulation is 6900 Å, a wavelength considerably less energetic than that equivalent to the forbidden-gap width. HALSTED *et al.*[18] incidentally have observed edge emission excited by long-wavelengths which they attribute to a two-step excitation process involving copper.

Stimulation at any wavelength shorter than this threshold produces a large change in conductivity at nitrogen temperature. When stimulated with light, a crystal whose dark conductivity is 3×10^{-13} mho-cm^{-1} increases to a saturation conductivity of approximately 3×10^{-3} mho-cm^{-1}. The change in conductivity is slow and integrates over a period of time when illuminated by a *H*100*SP*4 mercury light with a Kopp No. 41 filter.[19] A sample illuminated in this fashion has been found to maintain its high conductivity state in the dark and at nitrogen temperature without further stimulation and without noticeable drop in conductivity for periods of at least 66 hr.

Most interestingly, optical irradiation is not required for stimulation of the crystal. If a crystal in the unstimulated, or low-conductivity state is removed from the nitrogen bath and allowed to warm in the dark to a temperature above approximately −80°C and is then returned to liquid nitrogen temperature, it is found possible to excite edge emission by mechanical tapping. The characteristics observed are now somewhat different, in that only the gentlest of tapping, sometimes even the slight rattling due to boiling of the nitrogen, is sufficient to excite emission. Furthermore, the flashes observed are not as localized as those obtained in the optically-stimulated case, but seem to occur more uniformly throughout the body of the crystal. Also, only three or four flashes are usually obtainable before exhaustion of the effect occurs. Fig. 1 is a photograph of thermally-stimulated emission produced in a rectangular crystal at nitrogen temperature.

In either case, pulses of green light are obtainable only around nitrogen temperature and possibly lower, in agreement with the observation that normal green emission disappears at higher temperatures. Warming the sample above liquid nitrogen temperature eventually causes thermal stimulation of the mechanical effect, but before this threshold is reached, it is found that pulsed output can no longer be observed in a sample stimulated either thermally or optically, i.e. the temperature required for thermal stimulation is above that at which edge emission disappears. The pulse shapes have been observed in cases of both types of stimulation. With optical stimulation, the pulse width is of the order of 100 μ in duration, but the exact shape and width are strong functions of the nature of the individual impact. The thermal stimulation, on the other hand, gives rise to pulses which are quite regular in shape, as Fig. 2 shows, where the rise time is approximately 0·2 μ sec and the time for decay to half-maximum value is about 0·8 μ sec. Measurement of the light output intensity with a calibrated photomultiplier indicates that about 10^8 photons are emitted in one flash from a thermally-stimulated crystal.

If a crystal of the type used in these experiments is stimulated with short-wave length radiation while at liquid nitrogen temperature and is then excited with infrared, edge emission occurs. The long wavelength threshold for excitation is 1·62 μ (0·76 eV), which will later be shown to be consistent with our model for mechanically-excited luminescence as reported earlier.[20]

The properties of crystals which have been specifically doped with impurities are of particular interest. Single crystals of CdS:Ag exhibit green emission under ultraviolet irradiation at liquid nitrogen temperature which masks a red component of emission at 6400 Å.[21, 22] At higher temperatures, green edge emission no longer occurs, and the red luminescence may be observed. This radiation is polarized perpendicular to the *c*-axis,[23] in accord with the band model discussed earlier.

If a CdS:Ag crystal is stimulated either optically or thermally, mechanical stimulation is found to excite not only green edge emission, but also red emission which is masked by the former. This effect is in accord with the interpretation of the position of the Ag impurity presented earlier[7] and shortly to be discussed further here.

Copper as an impurity, however, presents results which are in agreement with our model but in potential conflict with the symmetries assigned to the band edges by BIRMAN[8]. We assume that copper is an acceptor in cadmium sulfide because of the sign obtained in Hall and thermoelectric investigations and because of the production of

p–n junctions in solar cells by copper diffusion. This assignment is supported by the work of BUBE and THOMSEN[24].

If the ordinary emission of CdS:Cu is examined, not at liquid nitrogen temperature where green luminescence obscures the red due to the copper, but at dry ice temperature, it is found that the latter is polarized not perpendicular to the c-axis,

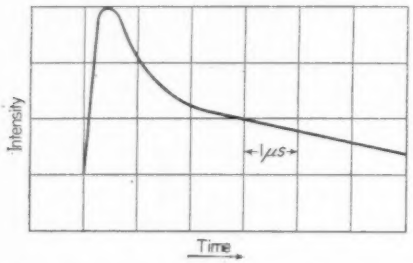


FIG. 2. Pulse shape of thermally-stimulated emission due to tapping.

as in CdS:Ag, but parallel to it.[23] Furthermore, increase in temperature decreases the polarization until only a slight component of polarized radiation is left at room temperature. If the assignment of symmetry to the conduction band and the levels near it is correct, then the copper level cannot have the Γ_9 symmetry of the top of the valence band, as might be expected. Furthermore, the depolarization of the copper emission allows the possibility of more than one copper acceptor level. Contrary to KLICK's statement[22] on the similarity of copper and silver levels in the sulfides, copper seems to have different properties from silver, at least in cadmium sulfide.

Unfortunately, the coupling of mechanical energy to the crystal states has been one of the most difficult problems to investigate. We have attempted to some extent to couple an ultrasonic transducer to the crystal, but it appears to date that shock, and not vibration alone, is responsible for releasing light. Experiments on calibrating the response due to impact are difficult to do in a quantitative way. The method of dropping spheres onto the surface from a known height was mentioned earlier, as was also the generation of waves by an ultrasonic transducer. In addition, we have attempted to produce piezoelectric oscillations in the crystals, but again we are hampered by the extremely high conductivity which occurs when the material is stimulated, as it must be for emission to occur. A similar restriction takes place in attempting to apply a high electric field to a crystal: the high conductivity in the stimulated state precludes the formation of an appreciable internal field, and this we take as evidence that triggering does not occur by production of a piezoelectric field by the shock.

Application of a field has been used to cause hole injection, and it has been found that such injection gives rise to edge emission as the recombination radiation. In addition, hole injection has been used to measure lifetime in a relatively pure crystal by application of a voltage pulse to an area contact of sputtered platinum and observation of the resultant green edge emission output with a photomultiplier and fast response circuitry. The initial results indicate a hole lifetime of the order of 1 μ sec in order-of-magnitude agreement with the time constant measurements of thermally stimulated mechanically-excited radiation described earlier.

This concludes the summary of experimental observations. In the next section we shall attempt to show that all these data are consistent with one particular model of the distribution of discrete levels.

INTERPRETATION

Let us first present a model for the band structure of the cadmium sulfide crystals dealt with here and then follow this with a discussion of the degree of agreement of the data presented with this model. Fig. 3 shows the level assignments for cadmium sulfide. The conduction and valence band have been represented with the multiplicity of valence bands and the symmetry assignments of the bands omitted for simplicity. These latter will be assumed to be those of BIRMAN *et al.* and should be kept in mind as being applicable here.

The uppermost level is assumed to be an electron trap level, taken to be 0·14 eV below the conduction band in accord with LAMBE *et al.*[2] Electron transitions are assumed to originate with this level. Next in order is silver, present only in CdS:Ag crystals. Although this is shown as one level at 0·6 eV,[24, 25] it is possible that two levels may exist here.[26]

Next in order is a hole trap level which is assumed to be near mid-gap (1·26 eV) at liquid nitrogen temperature. Below this is another hole trap level, approximately three-quarters of an electron volt from the valence band.

Finally, in copper-doped crystals, copper impurity sites are located approximately 0·83 eV energy-wise above the valence band. The energy

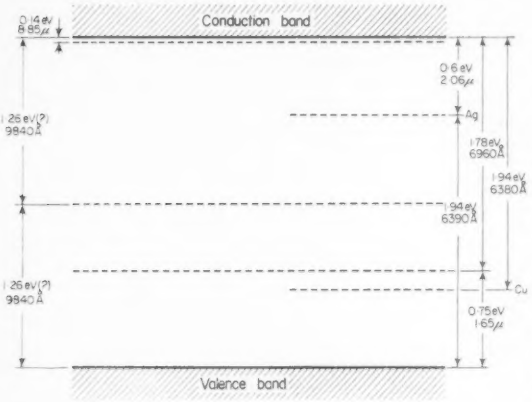


FIG. 3. Suggested population of forbidden gap in cadmium sulfide. Structure of valence and conduction bands has been omitted for simplicity. Question marks indicate tentative assignment of trap level. Silver and copper levels are to be assumed present only when crystal has been intentionally doped with either of these impurities.

and wavelength assignments are not meant to be exact, but only indicate the approximate positions of these levels.

Let us now consider the observations described earlier in the light of this picture. Optical stimulation of the sample can be accomplished by production of hole–electron pairs and subsequent trapping of the holes in one of the hole-trapping levels, or it can be achieved by stimulation of electrons from the lower hole trap level by wavelengths as long as 6900 Å. In the latter case, the holes are by definition "trapped" in this level. It is assumed, as in earlier literature,[1] that recombination of conduction electrons with trapped holes is forbidden, so that the conduction electron population gives rise to the persistent *n*-type conductivity and high gain observed in pure cadmium sulfide.

The lower set of hole traps is well-defined energywise. The depth of the level from the conduction band is determined by the fact that the threshold for optical stimulation of hole storage occurs at 6900 Å. The distance from the valence band should then be 0·75 eV to within approximately 0·01 eV at nitrogen temperature. That this is the case is shown by the fact that an optically-stimulated sample can be caused to emit radiation by illuminating it with infrared radiation. The threshold for excitation of the stored stimulation is 1·65 μ, which corresponds to the 0·75 eV activation energy predicted by the 6900 Å long-wavelength threshold for stimulation of the same level.

If the persistent conductivity produced by stimulation of this level is assumed to be due solely to conduction by a density of electrons equal to the saturation density of trapped holes, then a mobility of 10^2 cm^2/V sec yields a density of 2×10^{14} cm^{-3} for this level in the crystals showing this effect.

The situation for the upper set of hole traps is not as well defined. As mentioned earlier, a thermally-excited flash yields approximately 10^8 photons. Calculation of the level density which might give rise to this number of hole–electron pairs in a crystal of 1 cm^3 volume at room temperature indicates that a level 1·5 eV from the conduction band would have to have a density of 10^{22} cm^{-3}, which is patently impossible, whereas a level at mid-gap would require a density of 10^{18} cm^3. It is certain that this level must thus be at mid-gap or somewhat above, and we have arbitrarily shown it at mid-gap position in Fig. 3. In doped samples, the silver and copper levels are as shown in Fig. 3.

The polarization of the emission under ultraviolet excitation in Ag-doped samples is consistent with the light-emitting transition being from the silver level to the valence band, since the emitted radiation is perpendicular to the *c*-axis, consistent with the symmetries described earlier. The emission from copper, however, is parallel to the *c*-axis, in contradiction to what would be expected if the copper is derived from the Γ_9 and Γ_7 levels of the valence band.

With this one notable exception, the model described fits a large number of observations of mechanically-excited luminescence, conductivity, quenching, and the like, not only those presented

in this paper, but also as observed by other workers. It would thus appear to be a fairly good representation of the crystals showing mechanically-excited emission, although we have ample reason to believe it is incomplete.[13, 14, 17, 27]

One aspect of the problem of mechanical stimulation of emission which has not been fully treated is the nature of coupling between the lattice and luminescent centers. One reason is the difficulty inherent in investigating this coupling, although we have been attempting to find more quantitative ways of exciting the phenomenon. Our present speculation runs to considering the hole trap levels, which are deep and hence probably non-hydrogenic levels, as well-localized. It is then quite possible that little coupling exists between level and lattice and that a well-localized type of event, such as a dislocation passing by the level, or the vibration of a dislocation at which the level is situated, excites hole release. Such a description is consistent with the fact that the thermally-stimulated emission is much easier to excite, since it is derived from very deep levels.

It should be pointed out that such an obvious effect as coupling due to piezoelectric field must be discounted because the high conductivity concomitant with stimulation insures the absence of a field of any magnitude.

Acknowledgements—The authors wish to thank CHARLES H. BLAKEWOOD for his careful experimental work on this subject, and Profs. D. L. DEXTER and W. PAUL for helpful discussions.

REFERENCES

1. A review of the older literature which includes an extensive bibliography is that of BUBE R. H., *Proc. Inst. Radio Engrs, N.Y.* **43**, 1836 (1955).
2. LAMBE J. J., KLICK C. C. and DEXTER D. L., *Phys. Rev.* **103**, 1715 (1956).
3. FURLONG L. R., *Phys. Rev.* **95**, 1086 (1954).
4. DUTTON D., *Phys. Rev.* **112**, 785 (1958); DUTTON D., *J. Phys. Chem. Solids* **6**, 101 (1958).
5. LEMPICKI A., *Phys. Rev. Letters* **2**, 155 (1959).
6. KELLEY R. L. and FREDERICKS W. J., *Phys. Rev. Letters* **2**, 389 (1959).
7. LAMBE J. and KLICK C. C., *Phys. Rev.* **98**, 909 (1955).
8. BIRMAN J. L., *Phys. Rev. Letters* **2**, 157 (1959).
9. HOPFIELD J., *Bull. Amer. Phys. Soc.* (ser. 2) **3**, 409 (1958).
10. TAFT E. A. and HEBB M. H., *J. Opt. Soc. Amer.* **42**, 294 (1952).
11. WHEELER R. G., *Phys. Rev. Letters* **2**, 463 (1959).
12. COLLINS R. J., *Bull. Amer. Phys. Soc.* (ser. 2) **4**, 323 (1959).
13. DEXTER R. L., *J. Phys. Chem. Solids* **8**, 494 (1959).
14. SUGE Y., TANAKA S. and MASUMI T., Résumés des communications, pp. 79–80. Brussels Int. Conf. on Solid State Physics, June 1958.
15. GREENE L. C., REYNOLDS D. C., CZYZAK S. J. and BAKER W. M., *J. Chem. Phys.* **29**, 1375 (1958).
16. SIHVONEN V. T., BOYD D. R. and WOELKE C. D., *Phys. Rev.* **113**, 965 (1959).
17. REYNOLDS D. C., to be published.
18. HALSTED R. E., APPLE E. F. and PRENER J. S., *Phys. Rev. Letters* **2**, 420 (1959).
19. Note the sublinear relationship between emission and excitation intensity implying saturation of centers reported by DIEMER G., VAN GURP G. J. and MEYER H. J. G., *Physica s'Grav.* **23**, 987 (1957).
20. BLAKEWOOD C. H., WARSCHAUFR D. M. and REYNOLDS D. C., *Bull. Ame.. Phys. Soc.* (ser. 2); **4**, 45 (1959); WARSCHAUER D. M. and REYNOLDS D. C., *Bull. Am. Phys. Soc.* (ser. 2) **4**, 228 (1959); WARSCHAUER D. M. and REYNOLDS D. C., *Electrochem. Soc. Semiconductor Symposium*, 4 May 1959. See also infrared excitation and emission levels in phosphors given by BROWNE P. F., *J. Electronics* **2**, 1 (1956).
21. KLICK C. C., *J. Opt. Soc. Amer.* **41**, 816 (1951).
22. KLICK C. C., *Phys. Rev. Letters* **2**, 418 (1959).
23. WARSCHAUER D. M. and REYNOLDS D. C., *Phys. Rev. Letters*, **3**, 370 (1959).
24. BUBE R. H. and THOMSEN S. M., *J. Chem. Phys.* **23**, 15 (1955).
25. LAMBE J., *Phys. Rev.* **100**, 1586 (1955).
26. VAN GOOL W., *Philips Res. Rep.* **13**, 157 (1958).
27. BOER K. W. and GUTJAHR H., *Z. Phys.* **155**, 328 (1959).

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 13. pp. 257–270. Printed in Great Britain.

NUCLEAR MAGNETIC RESONANCE IN ALKALI ALLOY SYSTEMS, NaK AND NaRb*

L. RIMAI† and N. BLOEMBERGEN

Gordon McKay Laboratory, Harvard University, Cambridge, Massachusetts

(*Received* 6 *October* 1959; *revised* 16 *November* 1959)

VOL. 13 1960

Abstract—The Knight shift K in liquid NaK and NaRb systems is a linear function of the relative Na-concentration 1–C. The relative variation $K^{-1}(\partial K/\partial C) = 0{\cdot}361$ for Na^{23} in NaK, $= 0{\cdot}518$ for Na^{23} in NaRb and $= 0{\cdot}270$ for Rb^{87} in NaRb. An interpretation in terms of the scattering theory of conduction electrons is given. The line width of the Na^{23} and Rb^{87} resonances is a non-linear function of C. It is about a factor 5 larger in the Rb-rich alloys than at the Na-rich end. This result can be explained by the electron-coupled spin exchange interaction between different nuclei. The remarkable fact that these spin–spin interactions do not average to a small value in the liquid metals necessitates the postulate of a long range character for the electron-coupled interaction. The Knight shift and line width in solid terminal solutions have also been observed and compared with theory. The experiments show that nuclear magnetic resonance is particularly well suited to determine phase diagrams of alkali alloy systems.

1. INTRODUCTION

THE investigation of the electronic structure of alloys by means of nuclear magnetic resonance techniques[1] has been reviewed in a number of papers.[2–4] No measurements have been reported on alkali alloy systems. Yet these offer some unique advantages, which are evident in the corresponding case of the pure alkali metals.

In the first place the theory for the electronic structure of the alkali metals is most advanced. In particular, it is possible to calculate $\langle|\psi(0)|^2\rangle_F$, the density of the electrons at the position of the nucleus averaged over the Fermi surface with some precision.[5] BENEDEK and KUSHIDA[6] have checked the variation of this quantity with pressure in the alkali metals by measuring the Knight shift. A high precision is possible, because rapid self-diffusion causes a narrowing of the nuclear magnetic resonance lines.[7, 8] Similar narrowing occurs in other metals close to their melting point, but in the alkali metals the narrowing occurs in a convenient wide range of temperatures near room temperature.

A similar state of affairs should occur in solid and liquid alkali metal alloys. The theory of the scattering of electron wave functions in alkali alloys has been treated by ROTH[9]. The calculation of $\langle|\psi(0)|^2\rangle_F$ in alloys will be reviewed in Section 4. The theory of the electron-coupled nuclear spin interactions,[10–12] their motional narrowing and other causes for the line width will be discussed in Section 5. Section 2 gives a brief description of the experimental techniques.

The experimental results are presented and discussed in Section 3. The Knight shift, line width and relaxation time were measured between 0°C and 60°C as a function of composition in the NaK and NaRb system. These systems were chosen, because Li does not mix readily with the other alkali's and Cs presents severe chemical difficulties. The K^{39} resonance has not been observed, because of the smallness of its nuclear moment. Most data pertain to the Na^{23} and Rb^{87} resonance and a detailed comparison with theory will be made for these resonances. Occasional observations on the Rb^{87} resonance, with 73 per cent abundance but

* This paper is based on a thesis presented to Harvard University by L. RIMAI in partial fullfilment of the requirements for the degree of Doctor of Philosophy. The work was supported jointly by the Office of Naval Research, the Signal Corps and the Air Force Office of Scientific Research.

† Present address: Research Division, Raytheon Company, Waltham, Massachusetts.

smaller magnetic moment, also show agreement with theory.

2. EXPERIMENTAL TECHNIQUE

A Pound–Knight–Watkins spectrometer was used in all experiments.[13] The limited range of radiofrequency power in which this instrument operates turned out to be very satisfactory for the narrow lines in the liquid and solid alkali samples used. In the case of Na^{23} a fair degree of saturation could be achieved at the high end of the available power range. The sweep of the Pound oscillator frequency was as low as 20 c/s per minute to record the narrow lines accurately.

The magnetic field with a stability of 1 part in 10^5 was provided by a Varian 6″ magnet. Since frequencies had to be compared to at least this accuracy between samples of different concentrations and at different temperatures, the scheme developed by BENEDEK and KUSHIDA for the pressure dependence of the Knight shift was adopted. Details of the frequency measurement and stabilization of the magnetic field by a proton resonance driven by a frequency-stabilized oscillator, have been given in their paper.[6] A long term frequency stability of 1 part in 10^7 can be achieved, which is more than required for the present experiments. The adjustment for optimum modulation at the proton mineral oil sample is therefore not critical, and the modulation was always chosen as a fraction of alkali alloy line width. The magnetic field inhomogeneity over the sample was about 2 parts in 10^6 and caused additional broadening of the narrowest lines, observed in pure Na.

The sample holder and coil assembly were shielded by heavy brass walls which in turn were in intimate thermal contact with a brass container. Water flowed into this container from a large temperature regulated bath. The temperature in the sample holder could be varied from 0°C to 80°C and held constant to within 0·1°C.

The samples were prepared in vacuum. A proper mixture of the chlorides of two alkali metals, in proportion to the desired alloy composition, was inserted with an excess amount of metallic calcium in a soft steel crucible. This was inserted in a Pyrex tube, with several distillation chambers attached. Gradual heating of the steel crucible above the softening temperature[14] of the Pyrex (650°C) gives a complete reduction of the chlorides to the alkali metals which condense in the first chamber. Repeated tipping off and vacuum distillation gives the purified alloy in the final chamber which is part of the resonance sample tube. Octane which has been purified by reflux distillation in the presence of a large amount of liquid alkali metal is introduced in another arm of the vacuum system. It is cooled by liquid air during production of the alkali alloy. The octane is then allowed to warm up and evaporate through a tube, the walls of which have been coated with a fresh thin layer of alkali metal evaporated from an auxiliary side tube. The desired amount of purified octane is introduced in the sample tube with the alkali metals, which is subsequently tipped off. Heating of the Rb samples produced no visible bubbling in vacuum indicating the purity of the samples. Dispersions were made by ultrasonic vibration of the test tube heated above the melting point of the alkali metal. The particles had dimensions between 1 μ and 20 μ, remained shiny indefinitely and settled at the bottom of the test tube in a dispersion which contained about 50 per cent of metal by volume. The radiofrequency coil was wound around the bottom of the test tube, which had an outer diameter of 7 mm.

The particle size was small enough to eliminate skin depth effects. Their composition was reproducible to within 1 per cent as indicated by different experimental runs. The n.m.r. results of a NaK alloy prepared in a different manner by mixing known amounts of the two metals, indicate also that the composition corresponds to the ratio of the chlorides used in the preparation. Chemical analysis would not be more accurate than these tests and less appropriate, because it might include reaction products, such as oxides, not observed in the n.m.r. experiments.

3. EXPERIMENTAL RESULTS

(a) *Knight shift*

The $Na^{(23)}$ resonance frequency in NaK alloys is plotted in Fig. 1 as a function of concentration at two different temperatures. The frequency is measured with respect to the resonance frequency $\nu = 9532{\cdot}4$ kc/s in pure Na at 25°C. A few runs were made at a frequency of about 6200 kc/s and confirmed the expected proportionality of frequency shift and magnetic field. At 25°C there is a liquid phase for concentrations $C = 30$ per cent to 85 per cent potassium. The Knight shift is a

VOL. 13 1960

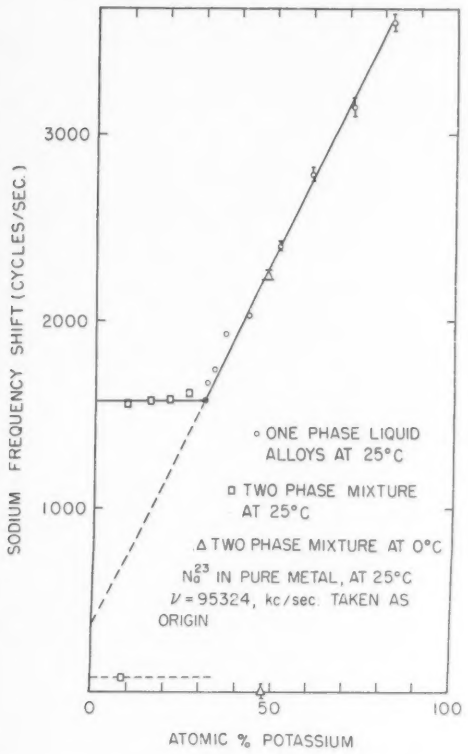


FIG. 1. The shift in the frequency of the Na^{23} resonance in Na–K alloys.

linear function of concentration. If the straight line is extrapolated to pure sodium, it cuts the axis at a positive intercept which corresponds very closely to the frequency difference of the resonances in pure solid and liquid sodium.[7] For K-concentrations, $C<30$ per cent, there is a two phase region with two resonances, which have intensities proportional to the relative concentrations of the two phases in the system. The resonance in the solid phase does not coincide with pure solid sodium. It indicates a solid solubility of K in Na. If one assumes the same slope of Knight shift vs. concentration in the solid terminal solution as in the liquid phase, one arrives at a solid solubility of $2{\cdot}5\pm0{\cdot}5$ per cent K in Na at room temperature. The few results at 0°C are also in agreement with the known phase diagram of the NaK system. This diagram was determined by cooling curves[15] and the terminal solid solubility was established by MACDONALD from electrical resistivity measurements.[16] The linear concentration dependence of the nuclear spin resonance frequency with concomitant sharp breaks at the phase boundaries suggests another way to determine phase diagrams.

The results in the NaRb system demonstrate the usefulness of this approach very clearly. Figures 2 and 3 show the Na^{23} and Rb^{87} resonance shift in NaRb alloys as a function of concentration at two temperatures. Note the choice of origin, appropriate to the pure solid metal in each case. The behavior of Na^{23} is similar to that encountered in the NaK alloys. The linear dependence on C in the liquid alloys extrapolates back to pure liquid sodium. There is a sharp break at the phase boundary.

The Rb^{85} and Rb^{87} resonance frequencies also increase linearly with decreasing Na concentration. The Rb resonances are lower than in pure Rb metal. There is a distinct temperature dependence

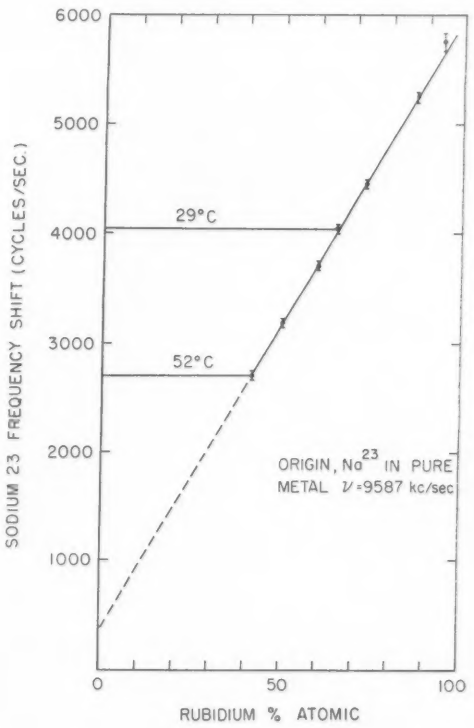


FIG. 2. The shift in the frequency of the Na^{23} resonance in Na–Rb alloys.

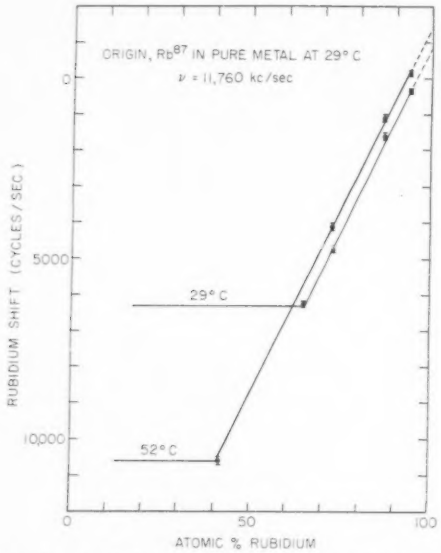


FIG. 3. The shift in the frequency of the Rb^{87} resonance in Na–Rb alloys.

of the shift, at a given concentration. At 52°C the straight line extrapolates to the value for pure liquid Rb at 52°C. At 29°C the extrapolation is to a frequency which corresponds to the resonance frequency of pure solid Rb at 29°C plus the discontinuity which occurs on melting.[7] In order to eliminate uncertainties in the composition of the alloys the Rb and Na resonant frequencies may be plotted against each other. A linear relationship results with a precision better than 2 per cent.

The temperature dependence of the Na^{23} resonance in NaRb alloys is shown in Figs. 4 and 5. As remarked above, there is no observable dependence in the liquid phase at constant C. If the liquidus in the phase diagram is met on cooling, C changes and the resonant frequency proportionally to it. The liquidus curve is thus traced out by NMR in Fig. 4.

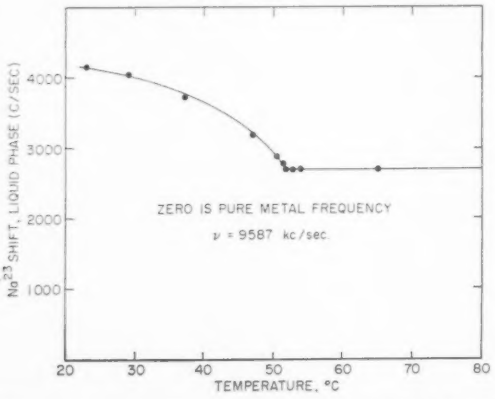


FIG. 4. The temperature dependence of the Na^{23} resonance in a sample containing 58 at.% Na, 42 at.% Rb. The curved portion traces the liquidus, on which the composition of the liquid phase varies.

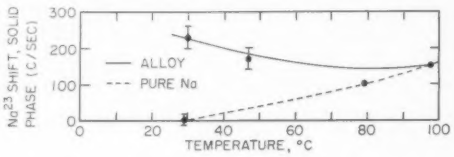


FIG. 5. The temperature dependence of the Na^{23} resonance in pure sodium and the saturated primary solid solution of Rb in Na.

The terminal solid solubility is determined by the difference in resonant frequencies of the solid phase and pure solid Na, shown as a function of temperature in Fig. 5. The phase boundaries may of course also be determined from the Rb resonance. This resonance has a larger explicit temperature dependence than Na as shown in Fig. 6. The extremely sharp break indicates the crossing with the liquidus. $(\partial K/\partial T)_C$ increases slowly with decreasing C, it is 30 per cent larger for a Rb concentration $C = 42$ per cent than for pure Rb.

The resulting points on the NaRb phase diagram are shown in Fig. 7. The liquidus had been determined previously from cooling curves.[15] The steady state n.m.r. method confirms the shape of the liquidus. The solidus has not previously been reported. n.m.r. establishes about the same solubility for rubidium as for potassium in sodium. The rubidium solubility is $3{\cdot}0\pm0{\cdot}5$ per cent at 29°C, and $2{\cdot}0\pm0{\cdot}5$ per cent at 47°C. The potassium solubility is $2{\cdot}5\pm0{\cdot}5$ per cent at 25°C.

Finally, the numerical values for the slope of the variation of the Knight shift with Na concentration $1-C$ in the liquid phase are given. $K_0^{-1}(\partial K/\partial C) = +0{\cdot}361$ for Na in NaK, $+0{\cdot}518$ for Na in NaRb and $+0{\cdot}270$ for Rb in NaRb. These values will be compared with theoretical calculations in Section 4.

VOL. 13 1960

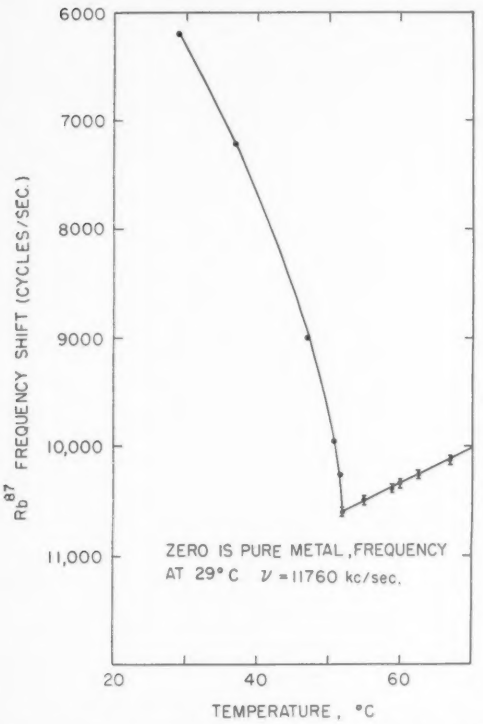


FIG. 6. The temperature dependence of the Rb^{87} resonance in the same sample used for the data of Fig. 5.

(b) *The line width and relaxation time*

The line width $(\Delta\nu)_{obs}$ is defined as the frequency separation between the extrema in the absorption derivative curve, recorded at the output of the phase-sensitive detector. The experimental line width is extrapolated to zero modulation amplitude. Lorentzian line shapes are expected for liquids and solid phases with rapid self-diffusion. Corrections for the inhomogeneity of the magnetic field ΔH are important especially for the narrow Na resonances in Na-rich alloys. The true transverse relaxation time is obtained from the observed width by

$$T_2^{-1} = \pi\sqrt{(3)}[\Delta\nu_{obs} - \gamma\Delta H/2\pi]$$

HOLCOMB and NORBERG determined T_2 for pure sodium directly, by pulse methods, $T_2 = 11{\cdot}8 \times 10^{-3}$ sec at 25°C. Thus $\gamma\Delta H/2\pi$ may be calculated from the observed width for pure Na to be 160 sec^{-1}. This correction is then applied to the other line width measurements.

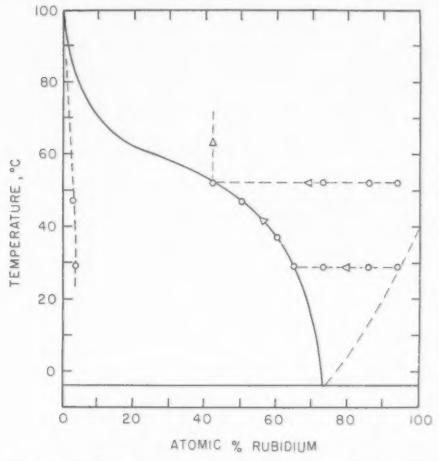


FIG. 7. The phase diagram for NaRb alloys. The drawn curve is the liquidus determined from cooling rate experiments. The experimental points for solidus and liquidus are determined from resonance data in Figs. 4–6.

The results for the inverse transverse relaxation time $(1/T_2)$ in NaRb alloys as a function of concentration are presented in Figs. 8 and 9 for the Na^{23} and Rb^{87} resonances, respectively. In addition the variation of T_1 for the Na^{23} resonance was determined from saturation curves. At the Rb-rich end the signal-to-noise ratio became poor

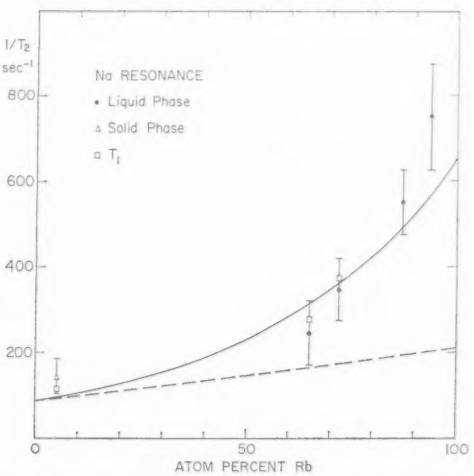


FIG. 8. The line width $1/T_2$ of the Na^{23} resonance in NaRb alloys. The width in the solid phase is the root mean square second moment $\langle\Delta\omega^2\rangle^{1/2}$. The width in the liquid phase is based on a Lorentzian shape.

and the available radiofrequency power was not sufficient to determine T_1. The three points obtained are shown in Fig. 8 and are consistent with the relation $T_1 = T_2$ within the wide margin (30 per cent) of error. The most striking feature is the rapid increase in width of the Rb-rich end. This increase of $(T_2)^{-1}$ with Rb concentration is much more rapid than would be expected from the Korringa relation[17] and the observed Knight shifts. This expectation is indicated by the dotted line. A detailed comparison with theory will be made in the next section. There is no detectable temperature dependence of the width in the limited range of investigation, and a decrease of 40 per cent in magnetic field did not reveal a field dependence of the line width. The data for the pure metals Na and Rb are not inconsistent with the results of HOLCOMB and NORBERG. Their data are more precise and encompass a much wider range of temperature and field. The present data are sufficient to reveal the remarkable dependence on concentration. This sheds further light on the T_1 and T_2 mechanisms, as will be discussed in the next paragraph. Some results for the Na^{23} line width in NaK alloys at 25°C were also obtained.

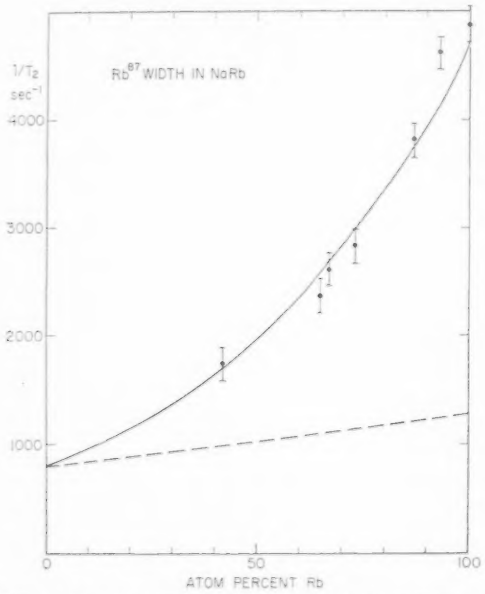


FIG. 9. The line width $1/T_2$ of the Rb^{87} resonance in NaRb alloys.

The value for $(1/T_2)$ in the terminal solid solution is 190 ± 60 sec^{-1}; in the liquid phase $1/T_2 = 80\pm40$ sec^{-1} and 220 ± 70 sec^{-1}, at 30 per cent and 71·5 per cent K, respectively.

The important features of the line width data may be summarized in the following way. (i) The line width of both the Na^{23} and Rb^{87} resonance increase rapidly in a non-linear fashion with Rb-concentration in Rb-rich NaRb alloys. This concentration dependence is in marked disagreement with KORRINGA's relation. (ii) The line width is independent of magnetic field within the experimental error. (iii) The line width is narrowed to some extent by diffusional motion, but the width in Rb-rich liquid samples is much larger than would be expected from a motionally narrowed interaction. (iv) The line width in sodium-rich solid primary solutions is noticeably broader than in pure sodium.

4. COMPARISON OF THE KNIGHT SHIFT WITH THE ELECTRON THEORY OF ALLOYS

The Knight shift $K = \Delta H/H_0 = \Delta\nu/\nu_0$ is proportional to the unbalanced electron spin density at the nucleus. When a solute atom is inserted in a matrix, the total charge density of the conduction electrons readjusts itself in the immediate vicinity of the solute atom, but remains unchanged outside a very small screening radius. The spin density important for the Knight shift, is however proportional to the charge density of only those electrons which have wave vectors on the Fermi surface.

BLANDIN and DANIEL[18–20] have considered the scattering of a Bloch wave $u(r)\exp(i\boldsymbol{k}\cdot\boldsymbol{r})$ by a spherical "square well" potential of radius r_s and depth ΔV. They obtain an expression for the relative change in charge density at a point $\boldsymbol{R}$ outside the scattering potential for electrons with wave number k_m, under the assumption that $u(r)$ is spherically symmetric and independent of $\boldsymbol{k}$. If variation of the Fermi level with concentration is negligible, the relative change in the Knight shift is

$$(\Delta K(\boldsymbol{R})/K_0) = \sum_l (2_l+1)[\{n_l^2(k_mR)-j_l^2(k_mR)\}\times \times\sin^2\delta'_l - j_l(k_mR)n_l(k_mR)\sin 2\delta_l] \qquad (1)$$

Here j_l and n_l are Bessel and Neumann functions of order l, and δ_l are phase shifts determined by

VOL. 13 1960

the boundary conditions on the surface of the potential well.

ROTH[21] has shown that the band structure of solvent and solute metal may be taken into account in a much more accurate manner. Her method gives a notable improvement in the calculation of the resistivity of alloys of noble metals and alkali metals. It may be expected that the improved scattering calculation will also lead to better values for the Knight shift. The general formulation of ROTH leads in the spherical band approximation to the same expression (1) of $\Delta K(\boldsymbol{R})/K_0$, except that the phase shifts are determined by a different boundary condition,

$$\tan\delta_l = \frac{kr_s j_l^2(kr_s)\left[kr_s\dfrac{j_l'(kr_s)}{j(kr_s)} - Rk'r_s\dfrac{j_l'(k'r_s)}{j_l(k'r_s)}\right]}{1-kr_s j_l(kr_s)n_l(kr_s)\left[kr_s\dfrac{j_l'(kr_s)}{j(krs)} - Rk'r_s\dfrac{j_l'(k'r_s)}{j_l(k'r_s)}\right]} \tag{2}$$

$R = f_p'(k'r_s)f_p(kr_s)/f_p(k'r_s)f_p'(kr_s)$, is the ratio of the logarithmic derivatives of the p-parts of the wave functions, for the solvent and the solute respectively, evaluated at the solvent atom radius r_s, at the Fermi surface.

$$u(\boldsymbol{r},\boldsymbol{k}) = f_s(\epsilon_F,\boldsymbol{r}) + i\boldsymbol{k}\times \\ \times \boldsymbol{r}[f_p(\epsilon_F,\boldsymbol{r}) - (r/r_s)f_s(\epsilon_F,\boldsymbol{r})] \tag{3}$$

For $R = 1$ the expression (2) reduces to the phase shift formula for plane waves used by BLANDIN and DANIEL, which involves the assumption that $u(\boldsymbol{r},\boldsymbol{k}) = f_s(\epsilon_F,\boldsymbol{r})$ is constant.

In the effective mass approximation the wave numbers at the Fermi surface are related by

$$\epsilon_0 + \hbar^2k^2/2m = \epsilon_0' + \hbar^2k'^2/2m' + \Delta V \tag{4}$$

Here ϵ_0 and ϵ_0' are the binding energies, m and m' the effective masses, which may be generalized to include a correction for the k^4 term at the Fermi surface. The potential difference ΔV is to be adjusted so that the phase shifts satisfy the Friedel sum rule with the correct polarization of the impurity cell. This charge is compensated by an equal and opposite charge in the screening potential outside the square well.[21]

The change in Knight shift given by equation (1) is large for nuclei which are nearest or next nearest neighbors of the impurity and decreases rapidly with alternating signs for large distances. For very low concentrations of the solute atoms it is sometimes possible to see a structure in the resonance line, corresponding to the different positions with respect to the solute atom. For concentrations C over a few per cent no fine structure is resolvable and the line should be described by its moments. The first moment or average shift is given by,

$$\langle \Delta K \rangle = C \sum_{\boldsymbol{R}} \Delta K(\boldsymbol{R}) \tag{5}$$

where the sum is over all lattice sites and independent scattering has been assumed.

The theory is strictly valid only in the limit of small solute concentrations. The effect of multiple scattering is neglected. If the scattering centers are far apart the incident wave on each scatterer may indeed be regarded as an unperturbed Bloch wave. One cannot reasonably expect the theory to be valid for solute concentration of over 10 per cent. The experimentally observed linear dependence, however, over the whole concentration range is very striking and appears to be valid for the Na^{23} resonance up to Rb concentrations of 95 per cent.

It is possible to extend the treatment of BLANDIN and DANIEL to the case that the resonance of a solute atom, present in very small concentration, is observed. This calculation applies to the Na^{23} resonance at the Rb-rich end. Consider first a single Na atom in a Rb matrix. The sodium Knight shift is now determined by the value of the wave function at the origin of the scattering center. The wave function inside the scattering potential has also been given by MOTT and MASSEY. Only the zero-order partial wave is non-vanishing at the origin. For an incident plane wave $\exp(i\boldsymbol{k}\cdot\boldsymbol{r})$ the value at the centre of a uniform potential well is given by

$$\psi(0) = k' \sin(kr_s+\delta_0)/k \sin k'r_s \\ = \cos(kr_s+\delta_0)/\cos k'r_s$$

where k' is the wave number inside the well and r_s is the radius of the solvent atomic cell; δ_0 is given by equation (2) for $l = 0$ and $R = 1$. The same expression also appears to be valid for the more general Bloch wave scattering, considered by ROTH. In this case $R \neq 1$ in equation (2) and k and k' should be evaluated with the appropriate effective masses. The relative change in the Knight shift with respect to the shift in the pure solute matrix is

$$\frac{\Delta K}{K_{\mathrm{Na}}} = \frac{K_{\mathrm{Na}}(r_s = r_{\mathrm{Rb}})}{K_{\mathrm{Na}}} \left\{ \frac{\cos^2(kr_s + \delta_0)}{\cos^2 k' r_s} - 1 \right\} \quad (6)$$

This equation gives the shift for Na in almost pure Rb. The first factor on the right-hand side contains the volume variation of the Knight shift. It can be calculated from the known volume variations[6] of $|\psi(0)|^2$ and χ_p and amounts to 1·17.

The initial slope $K^{-1}(\partial K/\partial C)$ at the Rb-rich end can also be estimated. Consider two Na nuclei in the Rb lattice. It is known from equation (2) how much a single Na nucleus changes the charge density at each Rb lattice site. The relative change in the charge density at the site of the second Na nucleus should be the same as a Rb nucleus would have had at that site. In almost pure Rb($C \sim 1$) this leads to the relation

$$K_{\mathrm{Rb}}^{-1}(\partial K/\partial C)_{\mathrm{Rb}} = \{K_{\mathrm{Na}} + K_{\mathrm{Na}}(C = 1)\}^{-1}(\partial K/\partial C)_{\mathrm{Na}} \quad (7)$$

The assumption is made that the Na-nuclei scatter independently. This is a crude approximation for Na nuclei on adjacent lattice sites. Substitution of experimental values into equation (7) gives $K_{\mathrm{Na}}^{-1}(\partial K/\partial C) = 0{\cdot}42$, at the Rb rich end, whereas the experimental slope is 0·51 at both ends. This should be considered as reasonable agreement.

The numerical results of equations (1–6) are unfortunately very sensitive to the choice of r_s and ΔV. DANIEL[20] uses BLATT's procedure to estimate ΔV. The impurity cell is assumed to have the same radius as the solvent matrix and an effective polarization charge density equal to the electronic charge times the relative change in the atomic volume of the impurity; ΔV is chosen so that the Friedel sum rule gives this effective charge. ROTH also takes r_s equal to the solvent atomic radius; ΔV is adjusted so that the polarization charge inside the impurity cell is exactly compensated by the screening charge outside the cell. The screening potential is assumed to be exponential with a Thomas–Fermi range q^{-1}.

We have not made any detailed assumption about the nature of the screening, but have calculated the Knight shift for a wide range of values of ΔV and thus of polarization. In our calculation partial waves up to $l = 6$ were retained and the summation in equation (1) was carried out to the second nearest neighbors in the *bcc* lattice. The summation over further neighbors was replaced by an integration. We find that with $R = 1{\cdot}09$, a maximum value for the Na^{23} Knight shift $K^{-1}(\partial K/\partial C) = 0{\cdot}2$ for a polarization of about $-0{\cdot}5$ of an electronic charge on a Rb impurity atom. This should be compared with DANIEL's value of $K^{-1}(\partial K/\partial C) = 0{\cdot}15$ for a NaRb alloy. He uses $R = 1$, an integration beyond the eight nearest neighbors, retains phase shifts up to $l = 3$ and determines ΔV with BLATT's procedure. Before a comparison with the experimental value, 0·51, is made, one should add 0·17 to the theoretical values to take account of the relative variation of χ_p and $|\psi(0)|^2$ due to volume changes on alloying. These changes are only approximately linear in C.

The theoretical model is qualitatively correct, but any quantitative agreement should be regarded as fortuitous. It is particularly serious that no choice of the polarization and screening gives the large observed positive value for the Knight shift. One would have to adopt a different value of r_s or a different arrangement for the neighboring atoms in the liquid alloy. The variation in Knight shift is more sensitive to the details of the scattering and screening potential than the resistivity. The latter is quadratic in δ_l whereas the Knight shift and the Friedel sum contain linear terms in δ_l. Since $R-1$ is very small compared to unity in the alkali alloys studied, the band structure effects are less important than the polarization effects. In fact, for zero polarization, one finds $K^{-1}(\partial K/\partial C) = 0{\cdot}06$ for Na^{23} in NaRb, which would make the discrepancy between theory and experiment three times worse than with optimum polarization.

The Na resonance in almost pure Rb exemplifies a situation which is unusual in scattering experiments. The Na spin is a probe for the scattered intensity inside the scattering potential. Equation

VOL. 13 1960

(6) yields, with $R = 0{\cdot}94$ for Na in almost pure Rb, a relative increase in the Na^{23} Knight shift at $C = 1$ which is too large by at least a factor 2 or has the wrong sign for a reasonable polarization on the Na atom.

The Rb^{87} resonance on the Na-rich end would have the wrong sign according to equation (6), if the same value of ΔV, with $R = 1{\cdot}09$ is used for which equation (1) gives the maximum value for $K^{-1}(\partial K/\partial C)$ of the Na^{23} resonance. The result of equation (6) is very sensitive to the choice of r_s near the point $k'r_s = \pi/2$. If ΔV is adjusted so that equation (6) gives the correct experimental value, then equation (1) gives an incompatible result.

The extrapolation of a straight line with a slope given by equation (1) near $C = 0$ across the phase diagram gives rather poor agreement with the prediction of equation (6) near $C = 1$.

This discussion also shows that the linear dependence of the Knight shift at all concentrations is perhaps a rather accidental property of the alkali alloys. In general, one should expect two different straight lines near $C = 0$ and $C = 1$, which should somehow be connected by a curved line at intermediate concentration.

The temperature dependence of the Knight shift has also been discussed by BENEDEK and KUSHIDA[6]. In the alloy one has

$$\left(\frac{\partial \nu}{\partial T}\right)_{C,p} = \left(\frac{\partial \nu}{\partial V}\right)_{C,T} \frac{\partial V}{\partial T} + \left(\frac{\partial \nu}{\partial T}\right)_{V}$$

The explicit temperature dependence $(\partial\nu/\partial T)_V$ is about the same for pure Rb and Na and is therefore expected to be independent of the concentration. $(\partial\nu/\partial V)_{Rb}$ is negative because of a dominant decrease of $\langle|\psi(0)|\rangle^2\epsilon_F$ with increasing volume; $(\partial\nu/\partial V)_{Na}$ is positive because the increase in density of states $\Omega n(\epsilon_F)$ at the Fermi level with increasing volume is dominant. The observed decrease of $(\partial\nu/\partial T)_{C,p}$ with Rb concentration C at atmospheric pressure may therefore be ascribed to the decrease in $(\partial\nu/\partial V)$ in going from pure Na to pure Rb.

The experimental data clearly rule out a liquid model with a large concentration of well-defined vacancies. Each vacancy would act as a scattering center with effective charge $Z = -1$. An enormous change in Knight shift and increase in resistivity would result in going from solid to liquid sodium, contrary to observations. Unfortunately, the increase in vacancy concentration deduced from diffusion and resistivity[22] just below the melting point cannot be confirmed by Knight shift measurements. The vacancy concentration is too low ($\sim 0{\cdot}1$ per cent) to give an appreciable effect on ΔK.

5. COMPARISON OF THE LINE WIDTH WITH THEORY

HOLCOMB and NORBERG[8] have given a detailed discussion of various factors determining the relaxation time and line width in pure alkali metals in the solid and liquid phase. Here the discussion will be extended to include the concentration dependence in alkali alloys. This will shed some additional light on some unexplained results in pure metals.

(a) *The interaction with the conduction electron spins*

The time average of this interaction determines the Knight shift. Since the correlation time for the time variation of this interaction is very fast ($\epsilon_F^{-1}h$), it leads to a contribution $T_2^{-1} = T_1^{-1}$ both in the solid and the liquid.[23] $(T_1)_e$, and $(T_2)_e$, have the same order of magnitude in the liquid and solid phase, since diffusional motion does not change the spectral distribution of this interaction which is determined by the electronic motion. It therefore is the dominant contribution to the width in pure liquid Na.

On alloying this contribution changes with concentration, since it is proportional to the square of Knight shift at each lattice site according to the Korringa relation. This is the correct procedure since the electronic motion is so much faster than the nuclear motion. The average relaxation rate becomes

$$(T_1)_e^{-1} = (T_1)^{-1}\left[1 + 2CK_0^{-1}\sum_R \Delta K(\boldsymbol{R}) + CK_0^{-2}\sum \Delta K^2(R) + C^2K_0^{-2}\sum_{R\neq R'} = \Delta K(\boldsymbol{R})\Delta K(\boldsymbol{R}')\right]$$

$$= (T_1)^{-1}_{\substack{\text{pure}\\ \text{metal}}}\left\{\left[1 + CK_0^{-1}\sum \Delta K(R)\right]^2 + C(1-C)K_0^{-2}\Delta K^2(R)\right\} \qquad (8)$$

The correction from the last term is of the same order as the multiple scattering effect.

Deviations from this relation are caused by correlation effects, but these do not depend strongly on C. They depend somewhat on the effective mass and atomic volume.

The hyperfine interaction between the nuclear spin and the "non-S" part of the electron wave function also contributes to T_1^{-1}, although it does not contribute to K. The variation of this contribution with C should be similar to the variation of the S-part, given by equation (8). The scattered intensity should be evaluated not at the center of the cell R, but an appropriate average over the "non-S" orbital in the atomic cell should be taken. One may conclude that the contribution $(T_2)_e^{-1}$ to the line width should exhibit the general quadratic dependence on C of equation (8). This contribution is shown by the dotted lines in Figs. (8) and (9).

In the rigid alloy lattice the dispersion of the Knight shifts contributes to the mean square deviation from the average position by an amount

$$\langle \delta\nu^2 \rangle_K = \nu_0^2 C(1-C) \sum_{\boldsymbol{R}} \{\Delta K(\boldsymbol{R})\}^2 \qquad (9)$$

with $\Delta K(\boldsymbol{R})$ determined by equation (1).

Again a random solid solution has been assumed and multiple scattering is ignored. This may cause some additional broadening around the 50 per cent composition. The order of magnitude of the second moment $\langle \delta\nu^2 \rangle_K$, which is proportional to the square of the magnetic field, should be equal to the square of the shift. For $\nu_0 = 10$ Mc/s, the data of Figs. (1) and (2) indicate that $\langle \delta\nu^2 \rangle_K^{1/2} \sim 1$ kc/s for a 50 per cent NaK or NaRb alloy. This is comparable to the dipolar width $\langle \delta\nu^2 \rangle_{\text{dip}}^{1/2} = 3$ kc/s in the Na rigid lattice. The experimental data are for liquid and solid samples with rapid self-diffusion. The two interactions will be narrowed in roughly the same proportion. TORREY[24] has given a general formulation which may be applied to any interaction whose dependence on the relative position $\boldsymbol{R}$ of two atoms is known. In particular it may be applied in the present problem to the interaction of a nuclear spin i with the average magnetization of the conduction electrons. This interaction is time dependent through the relative position of the solute atoms $\boldsymbol{R}_{ij}(t)$ with respect to the nuclear spin.

$$A(\boldsymbol{R}_{ij})\boldsymbol{I}_i \cdot \langle \boldsymbol{S} \rangle = h\nu_0\{K_0 + \sum_j \Delta K(\boldsymbol{R}_{ij})\} I_{zi}$$

Note that this mechanism contributes only to T_2, since only I_z occurs in the effective Hamiltonian. The transverse components of electron magnetization $\langle S_x \rangle = \langle S_y \rangle = 0$ average out with the very fast electronic motion.

In the case of rapid nuclear motion with correlation time, τ_c, the interaction with the z-component of magnetization, leads to an effective narrowing of the second moment (9) according to

$$(1/T_2)_K = 4\pi^2\nu_0^2 C(1-C) \sum_{\boldsymbol{R}} K^2(R)\tau_c \qquad (10)$$

There is no corresponding contribution to T_1^{-1}. The experimental data did not reveal a dependence of the width on magnetic field. The contribution from equation (10) is small compared to the dipolar contribution of equation (12) for the solid terminal solutions with $C = 3$ per cent. Therefore, it is unobservable in these alloys. For liquid alloys with $C = 50$ per cent, $(1/T_2)_K$ is not very small compared to the dipolar contribution $(1/T_2)_{\text{dip}}$, but both interactions are small compared to the other broadening mechanisms in the liquid.

It is tempting to attribute the observation by NORBERG and HOLCOMB that T_2 in pure liquid Li and Na is 10–30 per cent shorter than T_1 to this effect. The difference $(1/T_2)_{\text{obs}} - (1/T_1)_{\text{obs}}$ would then be identified with $(1/T_2)_K$. Although C cannot be identified with a vacancy concentration in the liquid, one would still have a dispersion of Knight shifts for the different short-range configurations in the liquid, which transform into one another with a characteristic time τ_c. The observed effect does not have the required proportionality to ν_0^2. With $\nu_0 = 10^7$ c/s, $\tau_c = 10^{-10}$ sec $\langle \Delta K^2 \rangle = 10^{-6}$ for pure liquid Na, equation (10) yields $(1/T_2)_K = 3$ c/s, whereas the observed effect is 11 times larger. Although this is certainly not a satisfactory agreement, the theoretical possibility for $T_1 > T_2$ in liquids is interesting. In Section 5(d) the possibility for an interaction in the liquid with a long correlation time will be discussed, which might give $T_1 > T_2$, independent of magnetic field.

(b) *Dipole–dipole interaction*

The second moment in the rigid lattice, averaged over a polycrystalline sample, with a random distribution of isotropic constituents p, is given by

VOL. 13 1960

the VAN VLECK[25] formula,

$$4\pi^2\langle\delta\nu_q^2\rangle_{\text{dip}} = \sum_p (1+\tfrac{5}{4}\delta_{pq})C_p\gamma_p^2\gamma_q^2\hbar^2 I_p \times$$

$$\times(I_p+1)(4/15)\sum_i r_i^{-6} \qquad (11)$$

In the case of the NaRb system one has $C_{Na} = 1-C$, $C_{85} = 0{\cdot}72C$, $C_{87} = 0{\cdot}28C$. In a solid with rapid self-diffusion or in a liquid $\langle\delta\nu^2\rangle_{\text{dip}}\,\tau_C^2 \ll 1$, the effective width is according to KUBO and TOMITA[26] or SOLOMON[27]

$$(1/T_1)_{q,\,\text{dip}} = (1/T_2)_{q,\,\text{dip}} = \tau_c\sum_p (1+\tfrac{1}{2}\delta_{pq})\times$$

$$\times c_p\gamma_p^2\gamma_q^2\hbar^2 I_p(I_p+1)(4/3)\sum r_i^{-6} \qquad (12)$$

In pure Na at 25°C the diffusion is sufficiently rapid for this last equation to apply. In fact, τ_c is so small that the dipolar relaxation is negligible compared to $(1/T_1)_e$. The same is true for the liquid phases.

The additional broadening of the Na resonance in solid solutions with 3 per cent K or Rb, nor the rapid increase in width of both Na and Rb resonance in the Rb-rich liquid alloys, can be explained by the electronic or dipole–dipole interaction. Two further mechanisms which are believed to be the cause of these line broadenings will now be discussed.

(c) *Quadruple interaction*

The electric field gradients around a solute atom in a rigid alloy lattice have been discussed elsewhere. The contribution of the quadrupole relaxation in the case of rapid motion[28] is

$$(T_1^{-1})_{\text{quad}} = (T_2^{-1})_{\text{quad}} = \frac{3}{40}(eQ/h)^2\times$$

$$\times\left\langle\left(\frac{\partial^2 V}{\partial Z^2}\right)^2\right\rangle_{av}\tau_c\frac{4I_q(I_q+1)-3}{I_q^2(2I_q-1)^2} \qquad (13)$$

The value of the electric field gradient at the position of a Na atom, which makes the quadrupole contribution equal to $(1/T_1)_e = 70$ c/s for the Na^{23} resonance, is found to be

$$\langle q_{zz}\rangle = e^{-1}\left\{\left\langle\left(\frac{\partial^2 V}{\partial Z^2}\right)^2\right\rangle_{av}\right\}^{1/2}(1+\gamma_\infty)^{-1}$$

$$= 2\times 10^{22}\text{cm}^{-3},$$

if one adopts $Q = 10^{-25}\text{cm}^2$, $\tau_c = 10^{-10}$ sec and a Sternheimer factor $(1+\gamma_\infty) = 5{\cdot}5$. This is a reasonable value for solid sodium with 3 per cent Rb or K.

One Rb atom in a next nearest neighbor site might be expected to strain the lattice sufficiently to give this order of magnitude for the gradient. The observed increase in width in the solid terminal Na solutions may therefore be ascribed to a quadrupole effect, although the electron coupled interactions discussed in the next section could also explain the additional broadening.

In liquid metals the r.m.s. value $\langle q_{zz}\rangle_{\text{r.m.s.}}$ probably has the same order of magnitude 10^{22}cm^{-3} due to instantaneous deviations from cubic symmetry in the liquid configurations. In alloys with about 50 per cent Rb $\langle q_{zz}\rangle_{\text{r.m.s.}}$ may perhaps be larger by a factor 3. The value of τ_c in the liquid is smaller than one used for solid diffusion by a factor 10. Therefore, the quadrupole broadening of Na, the liquid alloys should at most be the same in absolute value as in the solid terminal solution. The Rb quadrupole broadening may be larger by an order of magnitude because of the larger Sternheimer factor. Under no circumstance, however, could the quadrupole interaction decrease rapidly as Na is added to pure liquid rubidium. The observed broadening in liquid alloys, shown in Figs. 8 and 9, must therefore have another cause.

(d) *Indirect spin-exchange interaction*

A coupling of the form $A_{ij}\boldsymbol{I}_i\,.\,\boldsymbol{I}_j$ between nuclear spins via a super-exchange type mechanism with the conduction electrons is well established. (See Refs. 10–12). For the heavier isotopes this interaction can be considerably larger than the dipolar interaction. The interaction constant A is proportional to the product of the hyperfine interactions on the atoms i and j. If these interactions are assumed constant over a conduction band with spherical energy surfaces, A is proportional to the product of the Knight shifts of the nuclei i and j and may be written in the form

$$A_{ij} = K_iK_j\gamma_i\gamma_j\hbar^2\beta^{-2}2\pi E_F F(2k_mR_{ij}) \qquad (14)$$

with

$$F(x) = x^{-4}(x\cos x - \sin x) \qquad (15)$$

E_F is the Fermi energy and k_m the wave number at the Fermi surface. In an alloy the Knight shifts

are given by

$$K_i = K_0 + \sum_{\text{solute}} \Delta K(R_{ij})$$

There is therefore a correlation between the Knight shift K_i and the indirect spin-exchange coupling A_{ij} because the position of solute atom j occurs in both expressions. In order to keep the expression for the contribution to the second moment from the couplings A_{ij} relatively simple, K_i will be replaced by its average value. This should be a good approximation because usually many neighboring solute atoms will contribute to K_i. With this approximation the spin-exchange contribution to the second moment of the qth resonance averaged over the random solid solution is

$$\langle \delta\nu^2 \rangle_{\text{ex}} = \sum_{p \neq q} \tfrac{1}{3} C_p \gamma_p^2 \gamma_q^2 I_p (I_p \neq 1) \hbar^2 E_F^2 \beta^{-4} \left(\sum_R F^2(2k_m R) \right) \times$$
$$\times \left[K_q + (1 - C_q) \sum_R \Delta K_q(R) \right]^2 \left[K_p + C_q \sum_R \Delta K_p(R) \right]^2 \qquad (16)$$

The effect of motional narrowing may again be treated by TORREY's method, since the dependence of the spin-exchange interaction on the relative distance of a nuclear spin pair is known. If a correlation time τ_c' is adopted, the contribution to the line width in the case of rapid motion is given by[28]

$$(T_2)_{q,\text{ex}}^{-1} = 2\pi^2 \langle \delta\nu_q^2 \rangle_{\text{ex}} \tau_c' \qquad (17)$$

The correlation time τ_c' is not necessarily the same as τ_c for dipolar interaction.

The observed line width in solid rubidium at low temperatures leads to a second moment which is about 30 times larger than the classical dipolar formula equation (12) predicts. If one substitutes numerical values for E_F, k_m and the observed Knight shifts into equation (16), one finds an answer of the same order of magnitude as the dipolar second moment. It should be emphasized, however, that equation (16) is based on an oversimplified model of the band structure; it is a single spherical band approximation. A similar discrepancy has been observed between the calculated and observed spin exchange second moment in thallium. Therefore, the experimental value in the rigid lattice will be adopted

$$(1/T_2)_{\text{ex,solid}} = 2\pi \langle \Delta\nu^2 \rangle_{\text{ex}}^{1/2} \approx 10^4 \,\text{sec}^{-1}$$

It is remarkable that the observed line width in liquid rubidium is only a factor 2·3 times narrower. GUTOWSKI attributed this width to $(T_2)_e^{-1}$. The line width, however, does not increase proportional to the absolute temperature in the liquid phase. Furthermore, the observed variation with concentration does not follow at all the dotted line in Fig. 9, which represents the contribution of the conduction electrons. On the other hand, the concentration dependence is explained very well by equations (16) and (17) as shown by the solid lines in Figs. (8) and (9). The ratio of the Na^{23} to Rb^{87} width is also given correctly $(1/T_2)_{\text{ex, Rb}} = 7{\cdot}9(1/T_2)_{\text{ex, Na}}$ at the Rb-rich end. In the plot for the Rb^{87} resonance the $Rb^{87}-Na^{23}$ interaction has been neglected in comparison with the $Rb^{87}-Rb^{85}$ interaction. The difference in behavior of T_1 and T_2 for the two Rb isotopes in metallic rubidium observed by HOLCOMB and NORBERG can also be understood in terms of this interaction.

The main difficulty is that the magnitude of the observed width in liquid rubidium requires a very long correlation time for the electron coupled spin interaction. According to equations (16) and (17) a value $\tau_c' \approx 5 \times 10^{-5}$ sec is required.

In principle, the exchange interaction may have a much longer correlation time than the dipolar interaction, e.g. if molecular association occurs. It is not compatible with the properties of metallic liquids to assume that two atoms remain associated as neighbors for times longer than a microsecond. The hypothesis of large semi-rigid islands in the liquid phase would give about the same correlation time for both the dipolar and exchange interaction. The data on light metals, such as Li, Na and Al,

show conclusively the motional narrowing of dipolar interaction caused by self-diffusion. The broad lines in the liquid phase appear to be characteristic of the heavier metals such as Tl, Hg, Rb and Rb-rich alloys.

The electron-coupled nuclear spin–spin interactions which are dominant in the heavier metals can have a long range, and consequently a long correlation time, if the reduced scattering vector is small. Suppose that the Fermi level nearly touches the Brillouin zone boundary. Consider the virtual scattering from an occupied state with wave vector $\frac{1}{2}\boldsymbol{K}-\Delta\boldsymbol{k}$ to an unoccupied state $\pm(\frac{1}{2}\boldsymbol{K}-\Delta\boldsymbol{k}')$. Then the quantity $x=2k_mR$ in equation (15) has to be replaced by $x=|\Delta\boldsymbol{k}\pm\Delta\boldsymbol{k}'|\,R$ for these scattering processes near the zone boundary. A similar state of affairs would occur if a fraction of the conduction electrons in another band had a very small effective mass and consequently a small value k_m. Irregularities at the Fermi surface because of the proximity of the zone boundary are known to occur in Cu and several other metals. It is a somewhat unexpected but possible situation in the liquid alkali metals.

Since $F(x)\propto x^{-1}$ for $x\ll 1$, but $\propto x^{-3}$ for $x\gg 1$, and since a given nucleus will interact with many more nuclear spins for small $|\Delta\boldsymbol{k}|$, even a very small fraction of conduction electrons near the zone boundary could give a significant contribution to the second moments (16) and (17). In particular the correlation time τ_c' would be proportional to the square of the interaction range or to $(\Delta k)^{-2}$. It is possible that the long-range coupling would be favored in the liquid phase because of the irregular band structure. In the absence of detailed theoretical work on more complex band structures, a long-range electron-coupled interaction will be postulated in the liquid metals, in addition to the short-range dipolar and exchange interactions which are averaged out by the thermal motion. Further experimental work on T_1 and T_2 in liquid alloys containing heavy solvent and light solute nuclei is desirable. It seems clear that the simple equation (14) for a single parabolic band with spherical energy surfaces needs considerable corrections.

The observations by HOLCOMB and NORBERG on the T_1 and T_2 of Rb^{85} and Rb^{87} in pure rubidium can be qualitatively explained in terms of a long- and short-range electron-coupled interaction. A coupled decay of two spin systems via spin exchange and conduction electron interactions will occur. The experimental data only give a simple exponential average rate of decay. A complete theoretical analysis, analogous to a discussion[29] of relaxation in the *HF* molecule, must await more precise experimental determination of the various characteristic decay times.

The anomalous decrease in T_2 on melting in Li and Na might likewise be explained by this type of interaction, if one is willing to admit the solution of some heavy impurities in the liquid phase.

(e) *Other interactions*

The contributions to the second moment or line width in the liquid of the various effects discussed so far are additive. This fact has been used implicitly in the preceding discussion. The contributions $(T_1)_e^{-1}$ and $(T_2)_e^{-1}$ are independent of the others because there is no correlation between the motion of the conduction electrons and the Brownian motion or spin orientation of the nuclei. The other effects have the following dependence on direction of radius vector $Y_2^m(\nu_{ij},\phi_{ij})$ and nuclear spin orientation of $\boldsymbol{I}_j$. Inhomogeneity of Knight shift has no dependence on either, dipolar interaction depends on both, quadrupolar only on the first, spin-exchange only on the second. Since the expectation values of $\langle Y_2^m\rangle$ and any nuclear spin component $\langle I_{x,y,z}^1\rangle$ vanish, there is no cross-correlation.

This would not be true if some other interactions of minor importance were considered. The pseudo-dipolar interaction should be added algebraically to the classical dipolar interaction. Its effect is estimated to be small compared to the scalar spin exchange in Rb, where S-type hyperfine interaction is dominant. This type of interaction might, however, be large between Pb^{207} spins in metallic lead. It could explain the large width in solid and liquid lead, although only one isotopic species with non-vanishing spin $I=\frac{1}{2}$ is present.

The p-character of the wave functions is also responsible for an anisotropy of the Knight shift at lattice sites in a solid alloy. The correction to the second moment, which would interfere with quadrupole effects, is very small.

Acknowledgement—The authors gratefully acknowledge helpful discussions with Professor H. BROOKS, Professor J. FRIEDEL and Dr. L. ROTH.

REFERENCES

1. BLOEMBERGEN N. and ROWLAND T. J., *Acta Met.* **1**, 731 (1953).
2. BLOEMBERGEN N., *Report on the Bristol Conference on Defects in Crystalline Solids* p. 1. Physical Society, London (1955).
3. KNIGHT W. D., *Solid State Physics.* (Edited by F. SEITZ and D. TURNBULL) Vol. 2, 93, Academic Press, New York (1956).
4. ROWLAND T. J., *Progress in Metal Physics* Vol. 9. Pergamon Press, London, in press.
5. BROOKS H. and HAM F., *Phys. Rev.* **112**, 344 (1958).
6. BENEDEK G. B. and KUSHIDA T., *J. Phys. Chem. Solids* **5**, 241 (1958).
7. MACGARVEY B. R. and GUTOWSKI H. S., *J. Chem. Phys.* **21**, 2114 (1953).
8. HOLCOMB D. F. and NORBERG R. E., *Phys. Rev.* **98**, 1074 (1955).
9. ROTH L., Thesis, Harvard University (1956); *Cruft Tech. Rep. No.* 267 (1957) (unpublished).
10. RUDERMAN M. A. and KITTEL C., *Phys. Rev.* **96**, 99 (1954).
11. BLOEMBERGEN N. and ROWLAND T. J., *Phys. Rev.* **97**, 1679 (1955).
12. YOSIDA K., *Phys. Rev.* **106**, 893 (1957).
13. POUND R. V. and KNIGHT W. D., *Rev. Sci. Instrum.* **21**, 219 (1950).
14. STRONG J., *Procedures in Experimental Physics*, Chap. 13. Prentice Hall, New York (1938).
15. BOHM B. and KLEMM W., *Z. anorg. Chem.* **243**, 69 (1939).
16. MCDONALD D. K. C., PEARSON W. B. and TOWLE L. T., *Canad. J. Phys.* **34**, 389 (1956).
17. KORRINGA J., *Physica, s'Grav.* **16**, 601 (1950).
18. BLANDIN A., DANIEL E. and FRIEDEL J., *Phil. Mag.* **4**, 180 (1959).
19. BLANDIN A. and DANIEL E., *J. Phys. Chem. Solids* **10**, 126 (1959).*

* The authors are indebted for receiving a copy of this paper before publication.

20. DANIEL E., *J. Phys. Chem. Solids*, to be published.
21. See Ref. (9); *Phys. Rev.*, to be published.
22. MACDONALD D. K. C., *Report on the Bristol Conference on Crystalline Defects in Solids* p. 383. Physical Society, London (1955).
23. WANGSNESS R. K. and BLOCH F., *Phys. Rev.* **89**, 728 (1953).
24. TORREY H. C., *Phys. Rev.* **92**, 962 (1953); *ibid.* **96**, 690 (1954).
25. VAN VLECK J. H., *Phys. Rev.* **75**, 1168 (1948).
26. KUBO R. and TOMITA K., *J. Phys. Soc. Japan* **9**, 888 (1954).
27. SOLOMON I., *Phys. Rev.* **99**, 559 (1955).
28. ABRAGAM A. and POUND R. V., *Phys. Rev.* **92**, 943 (1955), see footnote on p. 955.
29. SOLOMON I. and BLOEMBERGEN N., *J. Chem. Phys.* **25**, 261 (1956).

FOREWORD FOR THE TEN PAPERS WHICH WERE PRESENTED AT THE MELLON INSTITUTE SYMPOSIUM

VOL.
13
1960

The following ten papers in this issue represent a selection from a series presented at the *Symposium on Defect Structure of Quartz and Glassy Silica* held at the Mellon Institute, in May 1957. The papers have been selected and revised in the light of what has happened since the Conference so as to present a fair view of the current status of the subject. In one instance, an abstract only is included because the material has been published *in extenso* in another journal, to which reference is made.

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 13. pp. 272–278. Printed in Great Britain.

INTRODUCTION—DEFECTS IN SILICAS*

G. J. DIENES†

Brookhaven National Laboratory, Upton, New York

(*Received* 27 *May* 1959)

Abstract—As an introduction to the field of defect structure of quartz and vitreous silica the nature of bonding in silicas is reviewed briefly and a comparison is made between the crystalline and amorphous structures. It is found instructive to compare the status of the field with that of the typical ionic crystals, such as the alkali halides. Defects in silicas are then discussed more specifically under the following headings: (i) types of defects expected in silicas, (ii) production of defects by nuclear radiations and (iii) comparison of defects in crystalline and amorphous silicas on the basis of optical studies.

The purpose of this introductory paper is to review briefly the status of our knowledge of the defect structure of silicas and to compare it with that of such extensively studied crystals as the alkali halides. This introduction is meant to serve as a point of departure for the Conference itself and the new results reported at this Conference are not discussed. Consequently, practically all the remarks given here will be amplified and many perhaps superseded by the papers to be presented.

BONDING IN SILICAS

The nature of the bonding in a solid is an important factor in determining the defect structure of the crystal. For example, the fact that the alkali halides are essentially completely ionic plays a central role in interpreting structure sensitive phenomena in these crystals. It is natural, therefore, to start out with a brief description of the bonding in silicas. Surprisingly little work has been done on this problem. The best evidence for the character of the bonds is derived from electron density maps as obtained from careful X-ray studies. PETERS has done such a study and the following discussion is based on his work.[1]

It is easiest to bring out the essential points by comparing the electron density map of quartz with that of a typical ionic crystal, sodium chloride. The sodium chloride pattern is shown in Fig. 1. The important feature of this pattern is that one

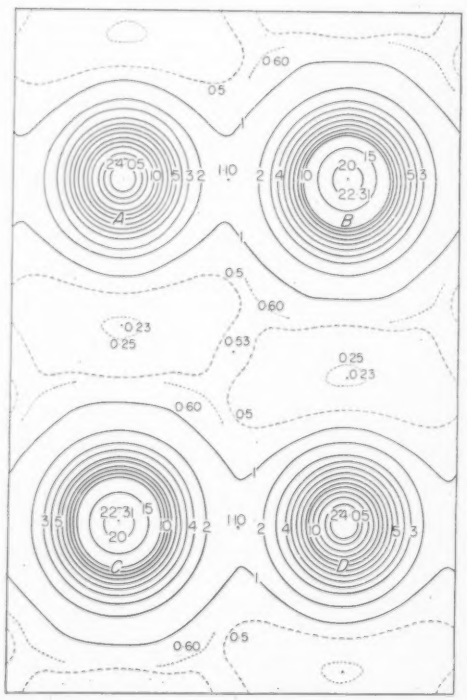


FIG. 1. Electron density map for NaCl in the [110] projection (from Ref. 1).

* Paper presented at *Symposium on Defect Structures of Quartz and Glassy Silica*, Mellon Institute, May 22–24, 1957.

† Under contract with the U.S. Atomic Energy Commission.

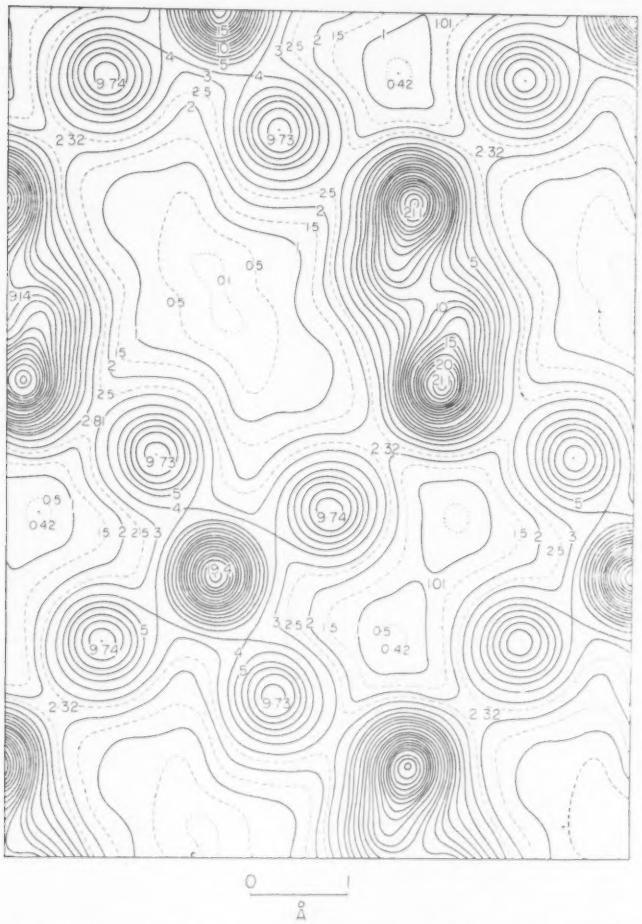


FIG. 2. Electron density map for quartz. Projection parallel to the two-fold axis (from Ref. 1).

finds large charge free regions in the Na^+–Cl^- direction (*A* to *C* in Fig. 1). This is clear indication of ionic bonding.

The quartz pattern is shown in Fig. 2 and is typical of mixed ionic and covalent bonding. The high electron density is clearly observable between silicon and the nearest four oxygens. A quantitative evaluation shows that the bonding is about half way between a purely ionic and a purely covalent bond, but somewhat closer to covalency.[1]

CRYSTALLINE AND FUSED SILICA

The detailed arrangement of the atoms is another important factor in determining the defect structure of solids. The structure of quartz is, of course, well known. Silica is, however, unusual among solids in that it is quite stable in the fused or glassy state. An important question is the relation between the crystalline and amorphous forms. From the classic work of ZACHARIASEN[2] and the X-ray work of WARREN and associates[3] one arrives at the conclusion that the local arrangement of the atoms in fused silica is very similar to that in quartz. Recent neutron diffraction work has not altered this picture.[4] The following points, based on the X-ray work, are particularly important:

(1) The average Si–O distance in vitreous silica is 1·62 Å in close agreement with 1·60 Å in crystalline silicates.

(2) The average number of oxygens around each

silicon is four in close analogy to the crystalline forms.

Thus, the amorphous material may be pictured as a random distribution of SiO_4 groups interlocked via oxygens which belong to two adjacent tetrahedra. From this picture a working hypothesis may be postulated, which many studies, like some reported at this Conference, will have to test in detail.

Properties which depend on local environment should be very similar in the crystal and the glass while those which depend on long range correlations are expected to differ greatly.

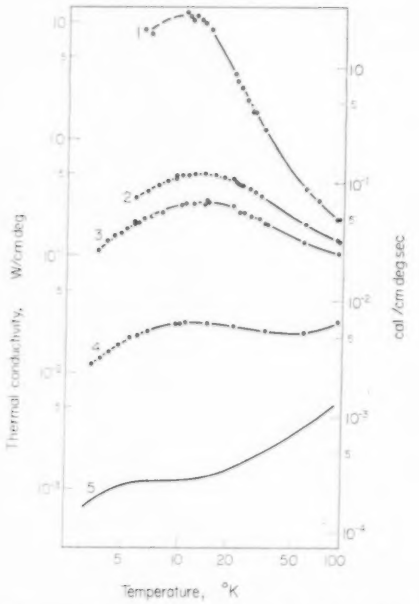


FIG. 3. Thermal conductivity of quartz crystal, irradiated quartz crystal, and quartz glass: (1) quartz crystal perpendicular to axis 5 mm square cross section: ● length measured 3·05 cm: ○ length measured 2·15 cm, (2) after first irradiation (1 unit), (3) ● after second irradiation (1·4 units); ○ after subsequent heating at 100°C for 3 weeks; (4) after third irradiation (16·5 units), (5) quartz glass (from Ref. 5).

It will be shown later that some of the color centers are indeed very similar in the two states. Thermal conductivities at low temperature, on the other hand, differ by orders of magnitude as shown in Fig. 3.[5] The figure also shows some data for fast neutron irradiated quartz. A drastic reduction in thermal conductivity is observed. The interpretation is rather intricate and will not be discussed here.[6]

COMPARISON WITH ALKALI HALIDES

In order to estimate the status of knowledge in the field of defect structures of silicas, it is of interest to compare it briefly to that of the alkali halides.[7] In general the comparison is very unfavorable to the silicas. Methods which have been employed for many years for studying the ionic crystals, are only coming into use, as this Conference clearly testifies. Some examples are given below.

(a) *Diffusion*

This important solid state process is well understood in simple ionic salts. There are many accurate investigations, the vacancy mechanism of diffusion is well established, activation energies for vacancy formation and mobility have been measured in many cases. As an example, a list of activation energies for mobility is given in Table 1.

Table 1. *Activation energies for motion of lattice vacancies, E_m, in alkali halides*

Substance	Vacancy	E_m(eV)
NaCl	Na^+	0·85
NaBr	Na^+	0·80
NaBr	Br^-	1·18
KCl	K^+	0·68
LiF	Li^+	0·65
LiCl	Li^+	0·41
LiBr	Li^+	0·39
LiI	Li^+	0·38

In contrast, the diffusion of only a few impurities has been studied in quartz and this by indirect methods.[8] A systematic study of the silicates has been started recently.[9]

(b) *Electrical conductivity*

A similar situation prevails for the electrical conductivity of silicates. The conductivity in alkali halides is known to be ionic, the conducting species have been identified in many systems and the vacancy mechanism of ion transport (via positive ion vacancies) is very well established for NaCl and KCl. Essentially nothing is known about

the fundamental conducting properties of silicas. Conduction by means of impurities has been measured in a few cases.[8]

(c) *Color centers*

The study of color centers has helped enormously in understanding the structure sensitive properties of alkali halides. Many of the color centers have been quite firmly identified with various defect structures as illustrated in Fig. 4. The various structures are identified in the caption. The

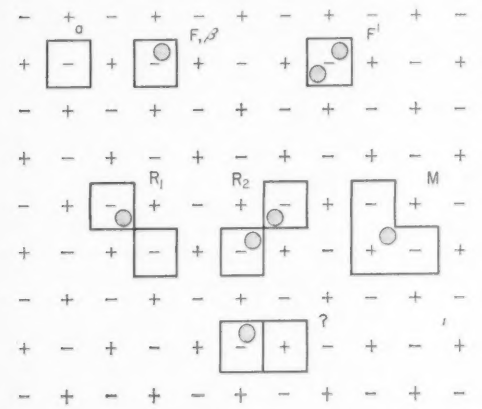


FIG. 4. Units believed to give prominent absorption bands in the pure alkali halides. Reading across and from the upper left-hand corner the units are: the negative ion vacancy presumed to be responsible for the α-band; the *F*-center consisting of the negative ion vacancy and an associated electron responsible for the *F*- and β-bands; the *F'*-center obtained by adding a second electron to the *F*-center; the R_1- and R_2-centers; the *M*-center; a hypothetical center consisting of a pair of vacancies and an associated electron. There is no experimental evidence for the last center—it probably decomposes spontaneously by ejecting a positive ion vacancy (from Ref. 7).

importance of such investigations has been recognized and this has been probably the most active field of basic research in silicas in the last few years. There are sufficient data at hand for a start at systematization but sound structural interpretation is a long way off.[10]

The obvious conclusion is that a tremendous amount of basic research work has to be done before one can understand the structure or defect sensitive properties of silicas. The problem is admittedly difficult. There are serious experimental difficulties because of the extremely high melting point of silica. There are difficulties of interpretation attributable to the complex crystal structure and bonding in these materials.

DEFECTS IN SILICA

Many physical properties of solids are sensitive to the presence of defects. Among the most important ones the following may be listed:

Optical (including radiation sensitivity),
Electrical (including dielectric loss),
Magnetic (particularly resonance),
Thermal conductivity,
Mechanical (including ultrasonics and internal friction, generally highly structure sensitive).

The relations between such physical properties and crystalline defects are intricate and are generally poorly understood even in the case of simple crystals. It should be emphasized that often one has to play a difficult "double game" in this field, i.e. use the various physical properties for studying defects and use our knowledge of defects in investigating the basic factors which govern the physical properties themselves.

As to the nature of defects expected in silicas one must rely on analogy to other solids since no theoretical calculations have been made from which the probability of occurrence, stability, mobility, etc., can be judged. The elementary defects are:

Impurities	Oxygen vacancies	Silicon vacancies
Dislocations	Oxygen interstitials	Silicon interstitials

The role of impurities is quite fully covered at this Conference. They are apparently of crucial importance in much of the color center work. The best known example is the substitutional aluminum center and its optical and electron resonance properties.[11]

One can say very little as yet about the role of dislocations in silica except that they are certainly present. Growth patterns typical of dislocations have been observed by WILLIS[12] and AMELINCKX[13].

A little more can be said about vacancies and interstitials from analogy with other solids. It is well known that thermodynamically such defects must be present in finite concentration (however minute) at any finite temperature. The reason is that the increase in the internal energy of the crystal is balanced by the entropy of mixing at

some finite concentration at any temperature above 0°K. Fig. 5 illustrates typical defect configurations in the alkali halides. For these materials calculations have shown quite conclusively that the energy of formation of interstitials is higher than that for vacancies, sufficiently so that

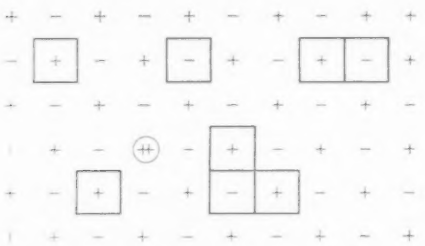


FIG. 5. Vacancies and vacancy clusters which play a prominent role in the theory of the alkali halides. As displayed from upper left to lower right and reading across, the units are: positive ion vacancy; negative ion vacancy; coupled pairs of vacancies of opposite sign; divalent ion present substitutionally with associated positive ion vacancy; cluster of two positive ion vacancies and one negative ion vacancy. It is believed that the coupled pairs may diffuse faster than isolated vacancies. It is also possible that the triple cluster diffuses rapidly (from Ref. 7).

they are present in completely insignificant amounts. Thus, only vacancies and their clusters are shown as identified in the caption. One might infer that the presence of interstitial oxygen is unlikely in stoichiometric SiO_2, although such a suggestion should be supported by detailed calculations.

Thermodynamically there are only three possible sets of point defects which may be formed in stoichiometric SiO_2. These are:

(1) Equivalent amounts of positive (missing Si) and negative (missing O) vacancies.

(2) Interstitial Si and an accompanying Si vacancy.

(3) Interstitial O and an accompanying O vacancy.

As mentioned before, interstitial oxygen may have a high energy of formation. The other defects are certainly expected to be present. Clusters of defects may also be formed, but it is not profitable to speculate about them at the present time.

No calculations have been made of electron redistribution around such defects. Again from analogy one would expect the following:

Si^i : traps electrons	O^i : traps holes
Si^v : traps holes	O^v : traps electrons

These considerations certainly serve as a useful guide but it should be emphasized that the situation is much more complex if the crystal is non-stoichiometric, i.e. contains impurities, excess silicon or excess oxygen.

EFFECT OF NUCLEAR RADIATION

The response of silica to various nuclear radiations has been of particular interest in the last few years and the subject is well covered at this Conference. Radiation has become an important tool for probing the defect structure of solids since rather large concentrations of defects can be introduced into a crystal this way. Defects normally present in insignificant amounts (say, interstitials in copper or sodium chloride) can also be introduced by radiation since thermodynamic limitations do not apply to their production by a highly irreversible high energy process.

At the present time the following mechanisms are considered important when radiation interacts with matter:[14]

(1) Displacement production by elastic collisions of fast particles with the atoms of the lattice. In SiO_2 the simplest result of such a process is the creation of vacancies and interstitials of both silicon and oxygen.

(2) Thermal or displacement spikes. The thermal agitation set up in the wake of a fast particle may rapidly heat a small volume of the material to very high temperatures, followed by rapid quenching. This process may result finally in highly disordered and in some cases vitrified regions in the solid.

(3) Ionizing radiation may produce three rather different effects:

(a) Scission and rearrangement of chemical bonds. This process is particularly important in organic materials but may well be of significance in SiO_2 since, as pointed out earlier, the bonds are about 50 per cent covalent.

(b) Production of color centers at existing defects by means of trapping of radiation released electrons.

(c) Production of defects by evaporation from dislocations or by means of the Varley mechanism. The latter involves multiple ionization of a negative ion and its subsequent displacement into an interstitial position where it becomes a neutral interstitial atom.[15]

One can find experimental support for all of these mechanisms and processes from studies of radiation induced property changes and their subsequent recovery in solids. In many cases alternative interpretations appear equally plausible at our present stage of knowledge. The response of SiO_2 to radiation is discussed in detail at this Conference and only some optical effects, which represent perhaps the most clear cut experiments, are discussed here very briefly.

VOL. 13 1960

SOME OPTICAL EFFECTS OF IRRADIATION IN SILICA

The recent work of MITCHELL and PAIGE[16] is discussed briefly because it throws light on the existence of point defects in SiO_2 and permits a rough comparison between the response of crystalline and amorphous silica to radiation.

It is quite well established that fast neutron irradiation is required to produce the 215 mμ (5·7 eV) and 160 mμ (7·6 eV) optical absorption bands in quartz at any appreciable rate. The 160 mμ band in particular is not induced by ionizing radiation at all. The 5·7 eV band can be bleached optically and is, therefore, associated with trapped electrons or holes. From the fact that irradiation with light in the 5·7 eV band bleaches this band and also the aluminum optical center (at 465 mμ), MITCHELL and PAIGE infer that an electron is being released from the 5·7 eV band and transferred to the Al center, and hence, that the 5·7 eV band is to be associated with a trapped electron.

From the results of heat treatment before irradiation, which indicated the presence of Si vacancies in unirradiated crystals, MITCHELL and PAIGE infer that the 7·6 eV band is not a Si vacancy since it could not be induced by ionizing radiation alone. Since there is rapid preferential bleaching of the 5·7 and 7·6 eV bands they conclude that the defects are either oxygen interstitial-vacancy pairs or silicon interstitial-vacancy pairs. Since the 7·6 eV band is not a silicon vacancy the remaining possibility is that the 5·7 eV band is to be identified with an oxygen vacancy which has trapped an electron and the 7·6 eV band with an oxygen interstitial which has trapped a hole. Much of this argument is inferential but these assignments may be adopted as a working hypothesis.

On the basis of the optical work on fast neutron irradiated silicas one may also say something about the crystalline vs. amorphous state. It was suggested earlier in this discussion that the local order would be very similar in crystalline and glassy SiO_2. There is some evidence for this from the optical work of MITCHELL and PAIGE. They found that the two prominent neutron induced bands discussed above are also present in fused silica and that they are only shifted very slightly in wavelength (5·7–5·8 eV, no shift in the 7·6 band). This is rather strong evidence for the similarity of the nearest neighbor environment. The two materials are not identical, however, since fused silica is more sensitive to ionizing radiation than quartz. Thus, defects are apparently produced with greater ease in the amorphous material.

SUMMARY

This introductory paper has given a rough sketch of the status of knowledge concerning the defect structure of silica. One can draw only one really firm conclusion. Research activity in this field has increased greatly in the last few years as evidenced by this Conference. There are very few concrete results as yet, but the importance of defects has been recognized and continued research in this field should certainly prove very rewarding.

REFERENCES

1. PETERS C., *Fortsch. Chem. Forsch.* **1**, 613 (1950).
2. ZACHARIASEN W. H., *J. Amer. Chem. Soc.* **54**, 3841 (1932).
3. WARREN B. E., *J. Appl. Phys.* **8**, 645 (1937); WARREN B. E., KRUTTER H. M. and MORNINGSTAR O., *J. Amer. Ceramic Soc.* **21**, 259 (1938) and other papers. For a recent review of the physics of the glassy state see CONDON E. U., *Amer. J. Phys.* **22**, 43, 132, 224, 310 (1954).
4. BREEN R. U., DELANEY R. M., PERSIANI P. U. and WEBER A. H., *Phys. Rev.* **105**, 517 (1957).
5. BERMAN R., *Proc. Roy. Soc.* A**208**, 90 (1951).
6. KLEMENS P. G., *Proc. Roy. Soc.* A**208**, 108 (1951).
7. For a thorough review of the alkali halides see SEITZ F., *Rev. Mod. Phys.* **26**, 7 (1954).
8. GIBSON G. and VOGEL R. C., *J. Chem. Phys.* **18**, 490, 1094 (1950); VERHOOGEN J., *Amer. Min.* **37**, 637 (1952).

9. LINDNER R., *J. Chem. Phys.* **23**, 410 (1955).
10. See for example the review by DAVIES T. H., 4th Int. Congr. on Glass, Paris, 1956.
11. O'BRIEN M. C. M., *Proc. Roy. Soc.* A**231**, 404 (1955).
12. WILLIS B. T. M., *Nature, Lond.* **170**, 115 (1952).
13. AMELINCKX S., *Natuurwet. Tijdschr.* **33**, 169 (1951).
14. See for example the recent review: BROOKS H., *Ann. Rev. Nucl. Sci.* **6**, 215 (1956).
15. VARLEY J. H. O., *Nature, Lond.* **174**, 886 (1954); *J. Nucl. Energy* **1**, 130 (1954).
16. MITCHELL E. W. J. and PAIGE E. G. S., *Phil. Mag.* **1**, 1085 (1956).

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 13. pp. 279–286. Printed in Great Britain.

A REVIEW OF THE GROSS STRUCTURAL EFFECTS OF ENERGETIC ATOMIC PARTICLES ON VITREOUS AND CRYSTALLINE SILICA*

W. PRIMAK†

Argonne National Laboratory, Lemont, Illinois

(*Received* 27 *May* 1959)

VOL. 13 1960

Abstract—The course of the changes resulting from irradiation and also those found on post-irradiation heating are reviewed. A hypothesis of the atomic behavior which will account for the observations is given.

INTRODUCTION

THE eventual effect of atomic displacement on most of the silica phases is to produce gross structural changes. These changes are readily observed in the X-ray diffraction patterns and greatly affect all of the physical properties. They are thus to be distinguished from the much smaller effects of ionization which seem to result in the trapping of electrons at selected sites, probably mainly impurities and imperfections already present in the substances, and which are manifested in changes in the absorption spectra and some of the most sensitive physical properties.

The most extensive investigations of the effect of atomic displacement thus far reported for the forms of silica are for material exposed in nuclear reactors. The major effects result from the slowing down of the silicon and oxygen atoms scattered by the energetic neutrons present in the reactor. From the results found for other substances, it is to be expected that some differences in behavior will be found for other kinds of atomic particle irradiation. However, the effects reported for boron containing glasses exposed in nuclear reactors (where the irradiating species are helium and lithium ions), for silicates containing naturally radioactive species (where the irradiating species are alpha particles and recoiling radioactive atoms) and for ion bombardment (e.g. argon ions) indicate the general course of the changes possess many similarities.

DESCRIPTION OF THE EXPERIMENTS

Evidence for the existence and occurrence of radiation disordered quartz and silica resulting from natural radioactivity is incomplete. Some of the variations in the properties of natural quartz[1] may be the result of natural irradiation, but it would be necessary to resolve them from the effects of variations in chemical composition and imperfect crystallization. Irreversible changes in the elastic modulus of natural quartz (as demonstrated by a change in frequency of a piezoelectric oscillator plate) caused by heating may be the annealing of changes produced by ionization.[2,3] FRONDEL[4] found effects on the frequency and the absorption spectra of oscillator plates made from smoky quartz which were similar to the effects produced by X-rays on clear quartz except that the effects which existed in smoky quartz annealed out at somewhat higher temperatures. The different annealing behavior of the smoky quartz may result from disorder caused by atomic displacement, or alternatively may be caused by impurities. FRONDEL and HURLBUT[5] compare the density, refractive indices and inversion temperatures of clear and smoky quartz and discuss the effect of composition but not of radiation. JOHNSON and PEASE[6] discuss evidence for a radiation effect in smoky quartz. HURLEY and FAIRBAIRN[7] report

* This paper was written in lieu of the one[15] presented at *Symposium on Defect Structure of Quartz and Glassy Silica*, Mellon Institute, May 22–24, 1957.

† Based on work performed under the auspices of the U.S. Atomic Energy Commission.

having sought unsuccessfully for extensively disordered quartz. HURLEY[8] feels that (if it exists there) disordered quartz should be visible at high magnifications in portions of thin sections adjacent to radioactive inclusions and notes that it has not been observed; and he suggests that the instability or the recrystallization of the disordered quartz under geologic conditions may be responsible for its absence. However, the ranges of the particles involved, the relative rates at which they cause disordering, and the effect of the juxtaposition of undamaged quartz should also be considered in this connection. Effects from natural irradiation are well known[9] for some silicates and other compounds. Irradiation cracking of glass and silica vessels was noted by early investigators[10] of the radioactive elements. The statement commonly found, that the cracking results from embrittlement is unconfirmed. It is now known that the dimensions of the glass may alter on irradiation,[11] non-uniformly under the conditions described. The cracking is a brittle failure resulting from the strains which are developed, but no increase in brittleness need be involved. Similar reports of brittle failures in the case of silica and glass irradiated in nuclear reactors have often been traced to thermal, mechanical or internal stresses which developed, rather than to a change in the intrinsic mechanical properties of the silica substance.[12]

The recent investigations which are considered here may be said to have begun with the studies by BERMAN *et al.*[13,14] of quartz irradiated in British and Canadian reactors. They found changes in the low temperature thermal conductivity which were particularly pronounced in the liquid hydrogen region. Relative dosages given for these irradiations are only approximate,[15] and it is difficult to effect a comparison with results obtained in other investigations. Upon learning of these studies, PRIMAK and FUCHS[16] arranged to have silica and quartz samples irradiated in the reactors at Hanford. The irradiated vitreous silica showed increases in density and refractive index, and the irradiated quartz showed decreases in density, refractive index, rotatory power and birefringence. STANLEY SIEGEL[16] found an increase in the lattice constants of the irradiated quartz and noted a drastic change in the X-ray scattering. Further irradiations were conducted at Hanford, in the Chicago piles, at the MTR, in the EBR and in the LITR to obtain irradiation doses which would give the whole range of the phenomena.[15–19] Independently, WITTELS[20–22] irradiated quartz in the LITR, studied its X-ray diffraction, and also made density measurements. WITTELS and SHERRILL[23] obtained the first set of systematic density data for the dosage range for which they reported.

Further investigations soon followed these initial ones. They will not be described in order, but rather, some attempt will be made to group similar effects. The changes in lattice constants given in one report from Brookhaven[24] do not correspond to the results of other investigators. WOODLEY[25] gives the volume of vitreous silica ampoules which had been used for irradiation, but the shrinkage seems somewhat excessive and may not have been determined sufficiently accurately or else some unexplained effects were involved; further, the exposure data seem inconsistent with other Hanford results.

Changes in the elastic constants as indicated by the change in frequency of quartz oscillator plates and caused by ionizing radiation are well known and seem to saturate at a small change.[2,3] It is to be expected that some of the other properties of quartz undergo change concurrently, but except for the absorption spectra, they have not been noted. The mutual influence of ionizing radiation effects and displacement effects seem unexplored and will therefore be neglected in what follows. Several fragmentary results for the irradiation of oscillator plates in nuclear reactors have appeared in reports from the reactor sites. A published work is that of JOHNSON and PEASE[26] for small irradiation doses in BEPO, and it also gives the changes in lattice constants. MAYER[27] studied the changes of all the elastic constants found for quartz irradiated in the French reactors. Additional data is given by MAYER and GIGON[28] who also describe the behavior of a number of properties of quartz and vitreous silica at low doses: expansion, thermal coefficient of expansion, elastic constants and heat capacity as shown by differential thermal analysis. The expansions found were much the same as those reported by previous investigators. The thermal coefficients of expansion showed no significant changes at these doses. Vitreous silica irradiated[29] for 7×10^{18} showed a 1·4 per cent increase in

VOL. 13 1960

Young's modulus and a 3·8 per cent decrease in the rigidity modulus. For quartz irradiated 22×10^{18} it was also found that some of the elastic constants increased and some decreased. The peak of the differential thermal analysis curve associated with the α–β quartz transition was found to be shifted to a lower temperature by several degrees, and to be decreased in area. The peak was restored to the values for unirradiated quartz on a second heating cycle. More recent work of theirs[30] has shown that the differential thermal analysis curves do not give a proper indication of the changes in the α–β transition. WITTELS and SHERRILL[23] noted the disappearance of the α–β transition in more extensively irradiated quartz. This was also reported by PRIMAK *et al.*[19]

LUKESH[31] studied the changes in the X-ray diffraction pattern of vitreous silica resulting from extensive irradiation. SIMON[32,33] made a similar study of extensively irradiated quartz and extensively irradiated vitreous silica (these had reached nearly the same state), and he concluded that the radial distance 0–0, Si–2^dO, and those of the more distant neighbor atoms were decreased and showed a wider distribution in the irradiated substances than in vitreous silica. Similar conclusions may be reached from LUKESH's work.[15] SIMON[33] also noted that the changes which occur in the infra-red absorption spectrum are similar to those found for vitreous silica permanently compacted under very high pressure. SIMON[34] found little change in the room temperature thermal expansion of vitreous silica as a result of extensive irradiation. FONER-COHEN[35] found an increase of a factor 2 in the low temperature thermal conductivity of vitreous silica resulting from irradiation. WESTRUM[36] found small changes in the low temperature heat capacity of vitreous silica and quartz as a result of irradiation in C.P.-5 and with his co-workers[37] has made additional measurements for extensively irradiated vitreous silica and quartz. Changes in the magnetic susceptibility caused by neutron irradiation were found by MCCLELLAND and DONOGHUE[38]. STEVENS[39] measured the paramagnetic component of the magnetic susceptibility introduced by neutron irradiation and found it to rise quite steeply, to reach a maximum at about 3×10^{19}, and then gradually decline to about 1/3 of its maximum value by 2–3×10^{20}.

Disordering of tridymite and crystobalite exposed in nuclear reactors was observed by PRIMAK[18], and by WITTELS and SHERRILL[23] but has not been studied in detail. WITTELS[22] found coesite to be quite resistant to such disordering.

KOCH[40] noted the effect of ion bombardment on the reflectivity of glass. HINES[41–44] studied the phenomenon in detail. He measured the effect of ion bombardment (mainly nitrogen and argon in the energy range up to about 50 keV) on the reflectivity of glass, vitreous silica, and quartz and interpreted it as caused by a change in the refractive index of the surface layers resulting from radiation damage.

The process of heating or the effects of heating irradiated material have been indiscriminately referred to as annealing, irrespective of the nature of the changes induced or their direction, or whether performed during or after irradiation. The annealing effects considered here are the changes in the properties resulting from post-irradiation heating. Among these investigations are BERMAN's[14] of the thermal conductivity of quartz; PRIMAK's *et al.*[19] of the density, refractive index and rotatory power of quartz and vitreous silica; WITTELS and SHERRILL's[23] of the density and X-ray diffraction patterns of quartz; PRIMAK and SZYMANSKI's[45]; and SZYMANSKI's *et al.*[46] of the kinetics of the density changes of vitreous silica. These studies reported the annealing of samples of a particular irradiation dosage. PRIMAK[15,47,48] measured the annealing of the density, expansion, refractive index, rotatory power and birefringence of vitreous silica and quartz irradiated for a number of different dosages. Since the behavior of the properties varies with irradiation dosage and with the property studied, no attempts will be made to give the details found in the particular investigations, but the general nature of the phenomena will be summarized below.

KLEMENS[49] interpreted BERMAN's thermal conductivity data to indicate the neutron irradiation introduced some point scatterers and also some relatively large defects (estimated as containing about 100 atoms) into the quartz. PRIMAK[16] considered that the property changes found after PRIMAK and FUCHS' original irradiation at 5×10^{19} indicated a gross disordering of the crystal. WITTELS[20,21] considered the changes to be due to

the packing of displaced atoms in particular interstitial positions. On observing the density–dosage curves reported by WITTELS and SHERRILL[23], PRIMAK[50] suggested that their increasing slope indicated the onset of a thermal spike mechanism of disordering developing with the irradiation induced decrease in the thermal conductivity. KINCHIN and PEASE[51] considered this increase in slope to be caused by a breakdown of the crystal structure. KLEMENS[52] suggested a mechanism of breakdown resulting from microscopic strains introduced by the admixture of radiation formed regions of much larger specific volume. PRIMAK[15,48] found that his own density–dosage curves were not as steep as WITTELS and SHERRILL's[23] and could not be accounted for by the onset of a thermal spike mechanism of disordering. However, the dosage scale of the changes in vitreous silica did support a thermal spike mechanism for this substance, and the physical properties of quartz require that thermal spikes should also be severe during the whole of the irradiation of quartz. These last considerations resolved much of the controversy concerning the changes and account, too, for the differential behavior of the different properties during irradiation and annealing. They are presented below.

INTERPRETATION

As a result of the above investigations it has been clearly established that when quartz is irradiated with atomic particles, it gradually becomes disordered into an amorphous form. This amorphous form is somewhat different from vitreous silica as ordinarily obtained, but the end products of the irradiation of quartz and vitreous silica are in most cases indistinguishable.[53] The case of vitreous silica is discussed first because it seems to be the simpler.[54]

A rapid compaction, amounting to 2–3 per cent of the volume of the vitreous silica, results from the irradiations described here. This compaction is completed by a neutron dosage of about 2×10^{19}, and it is followed by a gradual small dilatation amounting to about 10 per cent of the original compaction occurring over the following 3×10^{20} n°/cm^2. The small dosage required to effect the original change is indicative of large regions of the substance being altered by each event; and a correlation of the change with the neutron dosage establishes that a volume containing about 40,000 atoms on the average must be affected as a result of a scattering event. This is indeed the effective size of the region which it is anticipated will be raised to a high temperature for a short time as the result of the dissipation of the energy conveyed to the scattered atom, and it will be referred to here as the thermal spike. The vitreous silica structure consists of a three dimensional network of silicon and oxygen ions bonded in slightly strained configurations. Within the small region of the thermal spike, a deformation occurs to compensate for the thermally increased interionic distances. Permanent deformation of the Si–O tetrahedra themselves is not possible, nor is gross diffusion. However, a co-operative deformation of the configurations of the Si–O tetrahedra reminiscent of the rapid inversions which occur in the silica structure is possible, and the resulting compacted structure is frozen in as the thermal spike chills. There is thus some improvement of the local order at the expense of the long range order (at least in particular regions), and it is the destruction of this slight ordering which accounts for the slow secondary dilatation. The dosage scale of the secondary dilatation shows that it proceeds by the mechanism which will be described for the disordering of the quartz. During the compaction of the vitreous silica, the range of the distribution of bond distances is increased, and this must represent some kind of a bond strain. In the uniform slow heating of annealing, the original balance between the local and the longer range order is restored; the strain is relieved, and the material returns to its original state.

The property changes which take place on irradiation of quartz with atomic particles show two important features. The first is that there is an initial rate of change which is relatively slow followed by a period of greater rate of change leading to eventual saturation of the changes. It is convenient therefore, to describe the events as occurring in three stages which are not discrete but which gradually merge into each other. In the early part of the first stage the aeolotropy of the physical property changes seems to accord with the anisotropy of the crystal, or to the aeolotropy of similar changes produced in other ways. Thus the radiation induced dilatation shows the same aeolotropy as the thermal dilatation. In stage II there is a

VOL. 13 1960

marked departure from this behavior. Thus, although the substance becomes rapidly isotropic in this stage, the aeolotropy of the expansions do not correspond to becoming isotropic; and, as another example, calculations of the X-ray density from the diffraction peak maxima do not correspond to the density measured directly.

Another set of distinctions of the several stages of irradiation arises from the annealing behavior. After the early portions of stage I, the material returns to its original state in a simple manner if the material is gradually heated to a sufficiently high temperature. After irradiation to the later portions of stage I a phenomenon termed[55] anti-annealing (i.e. a change in the direction produced by further irradiation) appears in some of the properties (notably the expansions and the density change), and a delayed annealing in the case of some of the other properties. From the work of WITTELS and SHERRILL[23], it would appear that in the later portions of stage II the result of annealing is to produce a cryptocrystalline mass which, under some conditions, may be porous. In stage III the anti-annealing has dominated the whole of the annealing range, and the material becomes completely amorphous: to a state indistinguishable from vitreous silica if the irradiation has been conducted long enough, to a state which is slightly different if the irradiation has been somewhat shorter. The statements by WITTELS and SHERRILL[23] to the effect that the completely disordered quartz will crystallize on annealing is shown by the experiments of PRIMAK[15] to be incorrect. It is evident that their material must have been taken from the later portions of stage II when the appearance of X-ray diffraction patterns is much like those of a vitreous substance, but when some of the quartz structure must still remain as shown by the presence of some rotatory power and birefringence oriented at least approximately along the original axes. As long as there exist nuclei which are not deformed by the onset of annealing, crystallization can occur at the higher temperatures; these are temperatures at which silica devitrifies. The crystallization must depend on the nature and disposition of the nuclei which remain, and the results of WITTELS and SHERRILL[23] are additional support for the hypothesis presented below, that in the later portions of stage II the nuclei are isolated and not in phase with the original structure, and that as the irradiation progresses these nuclei become smaller and increasingly misoriented.

A mathematical investigation[56] shows that in the case of quartz it is the high temperature thermal conductivity that determines the temperature and the duration of thermal spike; and since the high temperature thermal conductivity of quartz is nearly the same as that of vitreous silica, the thermal spikes must be equally severe in the case of quartz right from the beginning of the irradiation, contrary to some earlier opinions expressed by the present author,[50] and thus will not account for the induction period represented by the slow rate of property change with dosage found in stage I. Despite their severity, the thermal spikes do not produce a gross disordering of the quartz. This is not entirely surprising since it is well known that the silica melts are very viscous, indicating that strong binding persists even in the liquid state. Further, it was seen that the thermoelastic phenomena in the case of vitreous silica favor an improvement of the short range order in the thermal spike. How then does the disordering of quartz take place? Here it is necessary to consider the role of the displaced atoms which could be neglected in explaining the rapid initial changes in vitreous silica because the latter is disordered to begin with. What must happen in quartz is that within the thermal spike, the displaced atoms are accommodated in the structure by co-operative movements of the configurations of the Si–O tetrahedra. Then with the accumulation of disorder, whole groups of atoms develop local configurations which are out of phase and even considerably rotated with respect to the neighboring structure. A rapid alteration of the physical properties ensues, and the material soon becomes grossly disordered. However the substance does not become completely amorphous (as indicated by the disappearance of the last vestiges of aeolotropy of the physical properties) until dosages of the magnitude required to displace every atom of the solid. This dosage is to be contrasted with the dosage required to produce the initial compaction of vitreous silica, about 1/40,000 as great.

Attempts to explain the detailed behavior of the various properties have been made by WITTELS[22] and by PRIMAK[15]. A repetition of the many details which are accounted for by the hypothesis given here is not warranted, and only a few matters which

are of special interest or which have been points of controversy are mentioned. It is clear by analogy with the cases of graphite, diamond and silicon carbide that simple unbound interstitial atoms do not accumulate in any large amount in irradiated silica, for such unbound interstitial atoms would give large amounts of stored energy, and this is not found. Neither would such atoms be expected in the silica structure in which the binding possesses much ionic character. However, atoms which are not in normal lattice sites are formed, and in the early part of stage I there result a number of partially bound atoms and concomitantly some unpaired electrons which give a paramagnetic contribution to the magnetic susceptibility. These displaced partially bound atoms and the undisplaced atoms with some broken bonds seem to act very much like random disorder. They do not behave like concentrated disorder located in discrete regions. If they did behave in the latter fashion, the aeolotropy of the radiation induced expansion would be similar to that of an elastic dilatation, and not like that of a thermal dilatation. WITTELS[21,22] considered that the subsequent changes were related to the crowding of oxygen interstitials into the *c* axis "channels". There can be no question that some interstitial atoms are located in these "channels", because these "channels" constitute all of the space not within an Si–O tetrahedron in the crystal. However, this view neglects the fact that by the dosage such crowding might become important, it is necessary to consider the positions from which these oxygen atoms originated. By this dosage, a considerable amount of rebonding has already begun to take place, and it is no longer possible to think of the effects as one of oxygen atoms being transported into undamaged regions of the crystal; the whole volume of the crystal has already been swept through by the fast knock-ons more than once. This is shown not only by comparison with the dosage scale for the changes found in vitreous silica, but by WITTELS'[22] own observations of the appearance of X-ray diffraction due to an amorphous component and to STEVENS'[39] observation that the paramagnetic component of the magnetic susceptibility has begun to decline. Another effect which has not previously been considered may be of importance here. The ionic volume of partially bonded oxygen atoms would be considerably less than the ionic volume of completely bonded oxygen atoms, especially those in strained configurations. This volume difference is out of all proportion to the concomitant opposite volume change of the silicon ions. Thus, when partially bonded oxygen ions become re-bonded into the structure, some additional distortion should be produced which would contribute to the increased rate of disordering during stage II. It is in this stage that KLEMENS[52] postulated changes are impelled by large shear stresses arising from the presence of adjacent regions of greatly different specific volume. The possibility of sustaining large stresses in small regions is indicated very directly by the measurements of HINES[53] who prepared amorphous films of irradiated silica on the surface of quartz. Although he did find some evidence for some strain broadening in the X-ray patterns, WITTELS[22] did not find the proportions of amorphous material required by KLEMENS' mechanism. The mechanism proposed here does account for the actual behavior. The wake of a fast knock-on does not produce a region of complete disorder; only a few scattered disturbances remain there. At the high temperatures in the thermal spikes originally associated with these wakes, changes which relieve at least the larger stresses can occur, and most of the stresses then remaining must be associated largely with the cooling of the thermal spike. Thus the large stresses hypothesized by KLEMENS never develop. They are not required as the driving force for the disordering, but the presence of much smaller stresses may be involved in the mechanism.

There is a feature of the disordering which seems quite characteristic and which appears also in the cases of the metamictization of the zircon class of minerals and which may occur in the disordering of other anisotropic silicates also. Quartz originally has an axial ratio of 1·10. When it is completely disordered, it is amorphous, hence isotropic, and therefore possesses an axial ratio of unity. However, if the physical expansions along the original axes are measured during the course of the disordering, the apparent axial ratio calculated from these results saturates during stage II at a value about 1·07. This effect must be associated with the formation of regions of the crystal greatly misoriented with respect to the original anisotropy. The large difference found by WITTELS[22] for the X-ray densities computed from the X-ray

diffraction patterns and those measured directly could also be accounted for by this kind of inhomogeneity arising in the partially disordered quartz. The non-parallel behavior of different properties has been explained on a similar basis.[15]

In coesite, accommodation of the displaced atoms by the mechanism proposed here for quartz cannot take place and coesite proves to be very resistant to radiation disordering.[22] On the other hand, the mechanism proposed here for quartz can operate equally well in tridymite and cristoballite, and they are disordered as readily as quartz.[18,23]

Silicates are considerably more resistant to disordering than quartz.[20] Eventually, as a result of displacement and the action of thermal spikes, a segregation of the silica to form radiation disordered silica seems to occur, and there also develop misoriented regions of the crystal similar to those hypothesized here for the case of quartz. A particularly interesting example is that of the zircon class of metamict minerals. The metamictization is caused by the alpha particles and the recoiling decaying natural radioactive species present in solid solution.[57] The former, having energies of about 3–4 MeV, produce scattered displacements (in the nuclear reactor scattered displacements are formed in the scattering of low energy neutrons) while the latter have energies from 0·05 to 0·1 MeV, of the order of magnitude of the energetic atoms formed by neutron scattering in nuclear reactors. The damage–dosage curves[58] possess many of the characteristics of the disordering of quartz. There is the induction period during which scattered disorder is accumulated, the second stage in which a rapid destruction of the crystalline order occurs, and the final stage in which the material is rendered completely amorphous. Some segregation of the oxides may result in the formation of new compounds in the second stage,[58] but eventually these, too, seem to become part of the amorphous mass. The general course of the disordering in the case of these metamict minerals seems very similar to the course of the disordering of quartz by atomic particles, and hence the latter may be considered the simplest example of a metamictization.

Acknowledgement—I am indebted to MARGARET FIELDHOUSE for having made an exhaustive search of the journal and report literature, both domestic and foreign. The literature of the field has been confused not only by publication restrictions, but also by contributions of directed research associated with materials testing. Thus relevant items are to be found in the early reports of nearly every one of the reactor sites, in this country dating to the nineteen-forties, and in other countries more recently. Although some of these items are referenced here, most of them are not. These omissions were entirely the choice of the author's, not Miss FIELDHOUSE's. The presentation of the references is not intended to establish any priorities in this field; the order was chosen in the hopes of achieving a compact readable presentation.

REFERENCES

1. SOSMAN R. B., *The Properties of Silica*. Chemical Catalogue Co., New York (1927).
2. FRONDEL C., *Amer. Min.* **30**, 432 (1945).
3. SEIDL F., *Acta Phys. Austr.* **10**(3), 169 (1956). This review gives some of the early observations of the effects of ionizing radiation.
4. FRONDEL C., *Phys. Rev.* **69**, 543 (1946).
5. FRONDEL C. and HURLBUT C. S., Jr., *J. Chem. Phys.* **23**, 1215 (1955).
6. JOHNSON F. B. and PEASE R. S., *Phil. Mag.* **45**, 651 (1954).
7. HURLEY P. M. and FAIRBAIRN H. W., *Bull. Geol. Soc. Amer.* **64**, 669 (1953).
8. HURLEY P. M., private communication, June 1958.
9. PABST A., *Amer. Min.* **37**, 137 (1952).
10. RUTHERFORD E., CHADWICK J. and ELLIS C. D., *Radiations from Radioactive Substances* p. 187. Cambridge University Press (1951).
11. LUKESH J. S., report KAPL 1307 (March 1955). There are also other investigations in the report literature.
12. Cf. THOMPSON M. W., report AERE-M/R-2305 (1957). There are also fragmentary results given in the report literature, some of them dating back to 1944.
13. BERMAN R., KLEMENS P. G., SIMON F. E. and FRY T. M., *Nature, Lond.* **166**, 864 (1950).
14. BERMAN R., *Proc. Roy. Soc.* A**208**, 90 (1951).
15. PRIMAK W., *Phys. Rev.* **110,** 1240 (1958).
16. The irradiation was arranged early in 1951 and the samples were examined later in that year. The first data appear in report ANL-4761 issued February, 1952. A partial bibliography of the report references is given in report ANL-4952, March, 1953. Papers were given at internal conferences in October 1952 and in March 1953 (report CF-53-3-276). Detailed public reports were delayed until a correlation of the Hanford irradiation data could be declassified.
17. PRIMAK W., FUCHS L. H. and DAY P., *Phys. Rev.* **92**, 1064 (1953).
18. PRIMAK W., *Phys. Rev.* **95**, 837 (1954).
19. PRIMAK W., FUCHS L. H. and DAY P., *Bull. Amer. Ceram. Soc.* **33**(4), program p. 25 (1954); *J. Amer. Ceram. Soc.* **38**, 135 (1955).

VOL. 13 1960

20. The first reports of this work are given in reports ORNL-1429, ORNL-1506, ORNL-1606 and ORNL-1677. A post-deadline paper was presented at an American Physical Society meeting in the spring of 1953. Other papers were presented at the American Physical Society meeting, May 1953 (Washington, D.C.) and at *Gordon Res. Conf. on Solid State Reactions in Ceramics, July*, 1954. Some information is given in report ACR-2 and in the paper by CRAWFORD J. C. and WITTELS M. C., *Proceedings of The International Conference on the Peaceful Uses of Atomic Energy*, Vol. VII, United Nations, New York (1956).
21. WITTELS M. C., *Phys. Rev.* **89**, 657 (1953).
22. WITTELS M. C., *Phil. Mag.* **2**, 1445 (1957) is a comprehensive paper which includes all of the material presented at this conference.
23. WITTELS M. and SHERRILL F. A., *Phys. Rev.* **93**, 1117 (1954).
24. KAMMERER O. F., report BNL-236 p. 33 (1953); also BNL-297.
25. WOODLEY R. E., report HW-40142, p. 31 (1955).
26. JOHNSON F. B. and PEASE R. S., *Phil. Mag.* **45**, 651 (1954).
27. MAYER G., *Action des Rayonnements de Grande Énergie sur les Solide*, p. 107. Gauthier-Villers, Paris (1956).
28. MAYER G. and GIGON J., *J. Phys. Radium* **18**, 109 (1956).
29. All of the neutron dosages in this article are given in terms of the damaging flux as defined by W. PRIMAK, *Nucl. Sci. Engng.* **2**, 320 (1957).
30. MAYER G., in a lecture given at the Argonne National Laboratory in November, 1957.
31. LUKESH J. S., *Phys. Rev.* **97**, 345 (1955).
32. SIMON I., *Phys. Rev.* **103**, 1587 (1956).
33. SIMON I., *J. Amer. Ceram. Soc.* **40**, 150 (1957).
34. SIMON I., *J. Amer. Ceram. Soc.* **41**, 116 (1958).
35. FONER-COHEN S., report ORNL-2413 (1957), p. 70.
36. WESTRUM E. F., Jr., *IVe Congr. Int. Verre, Paris*, 1956, *Proc.*, p. 1338.
37. To be published.
38. MCCLELLAND J. D. and DONOGHUE J. J., report NAA-SR-263; *J. Appl. Phys.* **24**, 963 (1953).
39. STEVENS D. K., *Phys. Rev.* **98**, 1541 (1955). See also report ORNL-2413 for a curve of the paramagnetic contribution as a function of irradiation dosage; this data was presented at this conference.
40. KOCH J., *Nature, Lond.* **164**, 19 (1949).
41. HINES R. L., *Phys. Rev.* **100**, 1267 (1955).
42. HINES R. L., *J. Appl. Phys.* **28**, 587 (1957).
43. HINES R. L., *Bull. Amer. Phys. Soc.* (ser. II) **2**, 18 (1957).
44. HINES R. L. described his work on the ion bombardment of vitreous silica and quartz and the annealing of the changes at this conference.
45. PRIMAK W. and SZYMANSKI H., *Phys. Rev.* **101**, 1268 (1956).
46. SZYMANSKI H., KIEFFER D. and PRIMAK W., unpublished. Presented at a Gordon Conference, 1957.
47. PRIMAK W., *Bull. Amer. Phys. Soc.* (ser. II) **1**, 136 (1956).
48. Data presented at this conference has been published elsewhere.[15]
49. KLEMENS P. G., *Proc. Roy. Soc.* **A208**, 108 (1951).
50. PRIMAK W., *Phys. Rev.* **98**, 1854 (1955).
51. KINCHIN G. H. and PEASE R. S., *Reports on Progress in Physics*. Vol. 18, p. 1. Physical Society, London (1955).
52. KLEMENS P. G., *Phil. Mag.* **1**, 938 (1956).
53. In HINES'[27] results an exception is found. The thin films of radiation-disordered silica formed on quartz and on vitreous silica did not have the same properties. This in part must have been the result of constraint; on the quartz base held in compression, and on the silica base held in a dilated state. However it is possible that an anisotropic amorphous state was in fact produced by the non-isotropic flux employed.
54. The mechanism hypothesized here was presented at this conference. It is described in detail elsewhere.[15]
55. HENNIG G. R. and HOVE J. E., *Proceedings of the International Conference on the Peaceful Uses of Atomic Energy*, Vol. VII, United Nations, New York (1956).
56. PRIMAK W., unpublished.
57. HOLLAND H. D., *Nuclear Geology (a Symposium)*, p. 178. Wiley, New York (1954).
58. HOLLAND H. D. and GOTTFRIED D., *Acta. Cryst., Camb.* **8**, 291 (1955).

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 13. pp. 287–295. Printed in Great Britain.

REACTOR AND GAMMA-RAY INDUCED COLORING OF CORNING FUSED SILICA*

P. W. LEVY†

Brookhaven National Laboratory, Upton, Long Island, New York

(*Received* 27 *May* 1959)

VOL. 13 1960

Abstract—When crystalline quartz, fused silica and many other substances are subject to reactor or gamma-ray radiations the samples develop optical absorption bands in the region 0·2–1 μ. Usually, there are many broad bands which overlap and resolution of the observed spectrum into the component bands is difficult. Corning fused silica is colored less than all of the materials studied and the observed spectrum can be resolved into its components by properly utilizing the variations in the relative absorption of the different bands created by changing the irradiation conditions. The most intense band is near 5·75 eV (218 mμ) and there is a smaller one at 5·05 eV (242 mμ) when the sample is irradiated in the reactor at 70°C. The 5·05 eV band has, relative to the 5·75 eV band, low intensity when the reactor irradiation is at 170°C but is much stronger when the irradiation is at liquid nitrogen temperature. Also, the peak of the 5·75 eV band shifts slightly with irradiation temperature. The band at 5·05 eV, which can be separated from the other bands without assuming a shape for it, is well fitted by a Gaussian curve and we have assumed that all other bands are similarly shaped. When a sample originally colored in the reactor is subsequently subjected to gamma-rays, additional absorption bands appear and their intensity is proportional to both the reactor and gamma-ray exposures. In all, new bands at approximately 5·5 (223 mμ), 4·5 (278 mμ), 2·0 (625 mμ) and one near 6·1 (200 mμ) eV, are present with indications of four others.

INTRODUCTION

For several years we have been interested in the reactor and gamma-ray induced coloring of crystals and other materials as a means of studying radiation engendered processes, particularly radiation damage. By this latter process we mean the disruption of a lattice by energetic particles such as the fast neutrons found in a reactor. Obviously the *F*-center coloring of the alkali halides is a measure of the number of vacancies these crystals contain and one may hope to find similar optical "handles" for all of the crystal imperfections, for example, absorption bands or glow peaks, to mention two of many possibilities, associated with interstitial atoms. Once one has succeeded in finding such "handles" for interstitials, vacancies or other imperfections, one can use them to determine the details of radiation damage phenomena.

During the course of surveying a number of likely materials to determine the extent of radiation induced coloring it was realized that a great many color centers were present in all of them and that it would be difficult to resolve the observed spectra into component absorption bands. The original measurements on Corning fused silica[1–3] showed that this material was colored less than the other substances tested and should be most susceptible to a detailed resolution into its component absorption bands. This was not entirely unexpected since most of the observed absorption bands are probably due to impurities. The other silica samples tested were all formed from natural crystal quartz which would be expected to contain more impurities than the Corning material which is made from purified $SiCl_4$. In this paper we discuss the resolution of the reactor and gamma-ray induced coloring of Corning fused silica into component absorption bands and give some preliminary results on how the formation of these bands is controlled by the various irradiation conditions.

* Paper presented at *Symposium on Defect Structure of Quartz and Glassy Silica*, Mellon Institute, May 22 and 24, 1957.

† Work performed under the auspices of the U.S. Atomic Energy Commission.

EXPERIMENTAL PROCEDURE AND RESULTS

All of the experiments described here used samples cut from a single melt of the Corning purified fused silica. The individual samples were slabs, 1–2 mm thick, with polished faces that were parallel to better than one per cent. Irradiations were either in the Brookhaven National Laboratory reactor or in a Co^{60} gamma-ray source. Samples exposed to gamma-rays could be measured immediately while reactor irradiated samples could be measured only after the induced radioactivity had decayed to a safe level. Optical absorption measurements were made with a Beckman DU spectrophotometer—with a few exceptions. Allowing the samples to stand in room light after irradiation caused optical bleaching of the order of one per cent. The bleaching that was observed occurred within a day or so after exposure to light and additional exposure did not produce additional bleaching. These comments refer to optical bleaching which must be separated from the effects of room temperature thermal bleaching that occurs whether or not the sample is protected from light. Room temperature thermal bleaching causes considerable changes which we will describe below.

Samples which were irradiated at above or below room temperature were all measured at room temperature (22°C). When samples were irradiated at liquid nitrogen temperature (LNT) they were kept in LN until they could be handled and then brought rapidly to room temperature. There are changes in the absorption spectra occurring within a few hours after the sample reaches room temperature which will be described in a later paper.

The influence of radiation conditions is apparent from Fig. 1, which illustrates this and several other features that will be discussed in turn. This figure was constructed in the following manner. The induced absorption spectrum was measured after reactor irradiations at 70° and 170°C and gamma-ray irradiations at 30°C, as a function of time in the reactor or of gamma-ray dose. It is found that the absorption at approximately 5·8 eV increases to a saturation value, which is difficult to determine, or to define precisely, and is very roughly the same for all of the reactor colored samples. Because a series of irradiations at LNT would be very protracted, the spectrum shown for −196°C resulted from a single irradiation. The gamma-ray induced

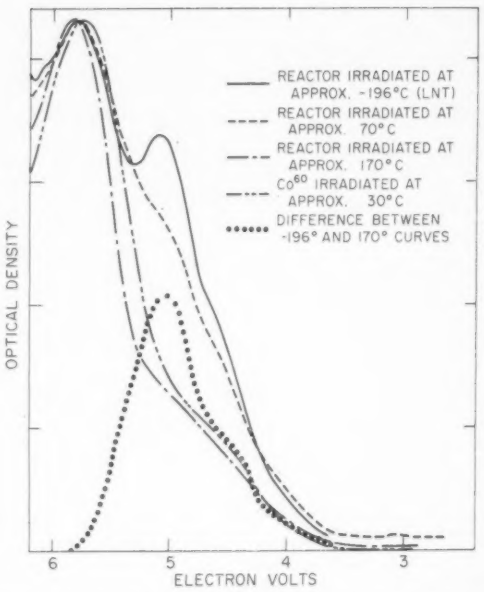


FIG. 1. The coloring induced in Corning fused silica by reactor irradiations at different temperatures. Also shown is the coloring induced by exposure to Co^{60} gamma-rays at approximately 30°C. The absorption spectra have been normalized to have the same height at the point of maximum absorbence which is slightly different for each case. The difference curve is shown in detail in Fig.2.

absorption appears to reach a saturation at roughly 1/3 the maximum absorption of the reactor colored samples. The four spectra shown represent coloring to approximately saturation and were normalized so that their maximum absorbance, near 5·8 eV, was the same. This maximum occurs at a different energy in each of the four spectra. Later it will be shown that more than one absorption band occurs in this region of the spectrum and that the relative intensity of these bands depends on whether gamma-ray or reactor irradiations are used and on the sample temperature during irradiation. Thus the position of the peak will be determined by the relative intensities of the bands that contribute to it.

The curves demonstrate that the intensity of the peak at 5·0 eV, relative to the peak at 5·8 eV, depends markedly on the irradiation temperature. They indicate that the 5·0 eV peak is larger, relative to the 5·8 band, the lower the irradiation temperature. It is important to emphasize that this

VOL. 13 1960

may not be the complete picture since we have obscured the fact that the spectra plotted all represent different amounts of radiation at slightly different irradiation rates as well as at different temperatures. Additional absorption bands are present near 4·5 eV, near 4 eV and at an energy greater than 6 eV, beyond the range of the spectrophotometer. The gamma-ray induced absorption spectrum is most similar to the absorption of the sample irradiated in the reactor at the highest temperature.

altered then equilibrium coloring is quickly attained. Also, assume that this equilibrium fraction is temperature independent. If the reactor radiations create the imperfection responsible for the 5·05 eV band it is probable that the imperfections would be more readily annealed out, i.e. destroyed, at a high temperature than at a low temperature. Thus in both cases greater coloring is expected at lower temperature. Not all of the assumptions can be expected to hold, and it is possible that both mechanisms can occur. Consequently, one may

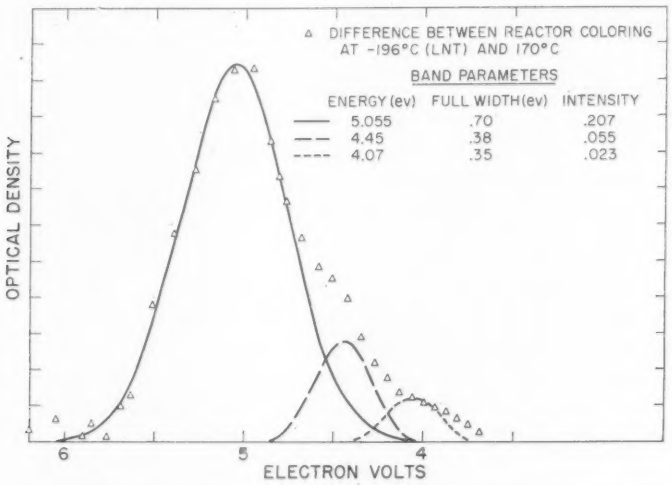


FIG. 2. The difference between the normalized curves of Fig. 1 that show the reactor coloring at −195°C and at 170°C. It is assumed that the data represents the three Gaussian shaped absorption bands shown. Obviously the largest peak is fitted quite well. An argument indicating why these bands should be Gaussian shaped is included in the text.

We will give two of many possible explanations why the 5·05 eV band would depend on the temperature of the sample during reactor irradiation. First, consider that the imperfection responsible for this band is not formed during the irradiation but that the coloring occurs because the center traps or loses an electron. If this process is temperature-dependent, so that it is less likely to occur or that it is less stable at high temperatures, then lowering the sample temperature would increase the number of stable centers, i.e. the optical density. For the second explanation assume that the equilibrium fraction of the colorable imperfections have been colored by the ionizing radiation present and that if the number of these imperfections is

ultimately find that this phenomenon is more complicated than we have considered here. Should the latter picture prove to be close to the actual situation we can readily understand why the gamma-ray induced spectrum is most like the spectrum obtained with a reactor irradiation at 170°C, the highest temperature used. During the gamma-ray irradiation only colorable centers, that existed prior to irradiation, can be colored, i.e. only a negligibly few colorable centers will be created. During reactor irradiation the equilibrium number of imperfections formed, which could be very small, depends on the ambient temperature. Then the coloring that occurs in both cases is due to imperfections that are stable, or have a fixed equilibrium

concentration, at the higher temperature. Centers due to impurities are a good example. It is entirely possible that the concentration of imperfections responsible for the coloring is determined during the formation of the silica.

To accomplish a resolution of the observed spectra into component bands we will assume that the same absorption bands are present in all of the samples and that the different irradiation conditions have merely changed the number of absorbing centers, i.e. the intensity of each absorption band. The difference in optical density between any two spectra should then indicate what bands are affected by changing irradiation conditions. The difference between irradiation at the highest and lowest temperature is shown in Fig. 1 and in detail in Fig. 2. It appears to consist of three separate bands.

In separating the observed data into the component bands it would be extremely convenient to know the shape of each band. A theoretical treatment of optical absorption in solids has recently been given by DEXTER[4] and part of his discussion deals with the theoretical shape of these bands. The classical Lorentzian and Gaussian shapes would appear to be special cases which would not be ordinarily encountered. A short argument can be given which supports the idea that in amorphous or glassy substances it is reasonable to expect absorption bands that can be approximated by Gaussian curves.

The shape of any absorption band should be Gaussian if it is determined primarily by the thermal fluctuations that influence the imperfection responsible for it. In an amorphous substance some of the regularities, such as crystal symmetry, that tend to cause deviations from the Gaussian shape do not occur. Thus one might expect there is greater probability that the absorption bands found in glassy substances will be Gaussian shaped than the bands found in crystals.

An ideal determination of the shape of an absorption band would require that it is not superimposed on other bands. At the moment we do not know of an isolated absorption band in fused silica that is intense enough to study carefully.

In Fig. 2 the largest band was first fitted by a Gaussian curve by simply picking values of the peak energy E_0, the maximum absorption α_m (intensity), and the full width at half maximum absorption U, in the equation

$$\alpha(E) = \alpha_m \exp\left[-\frac{4 \ln 2}{U^2} (E - E_0)^2 \right].$$

This is the ordinary equation of a Gaussian curve written in terms of the physical quantities we customarily use. This curve, with the parameters given, fits the experimental data quite well. The data in Fig. 1 and other data given below were subsequently analyzed *assuming that all the observed absorption bands were Gaussian.*

There are procedures, such as least-squares fitting, which can be used to systematically determine the parameters defining a Gaussian curve. When a sufficient amount of data has been accumulated, an attempt will be made to do this with a programmed calculator.

The three bands shown in Fig. 2 were obtained by fitting the largest peak by a Gaussian curve, with the parameters tabulated, to the data and plotting the difference between the solid curve and the data points. The resulting curve obviously contained at least two separate absorption bands which could be approximated by the curves shown. The fits are not perfect, but this procedure has made it possible to determine the parameters of at least one and perhaps two or three bands that are contained in a spectrum consisting of the superposition of at least four different bands. Obviously, the parameters given for the two smallest bands could be inaccurate.

The coloring produced by $2{\cdot}6 \times 10^8$ r of Co^{60} gamma-rays is shown in Fig. 3. In attempting to separate the observed data into component absorption bands two procedures were followed. The first of these assumed that there was a band at 5·055 eV and curves were obtained showing the difference between this band with different intensities (i.e. different values of α_m) and the observed data. By trial and error it was found that the data could be represented by three Gaussian shaped bands at 5·055, 5·76, 4·47 eV with an additional small contribution from a band at approximately 6 eV, at or below the operating range of the spectrophotometer. Some idea of how well this procedure works may be seen by noting in Fig. 3 the curve obtained when the bands at 5·055 and 5·76 eV are subtracted from the data points. The second procedure followed was to first fit a Gaussian

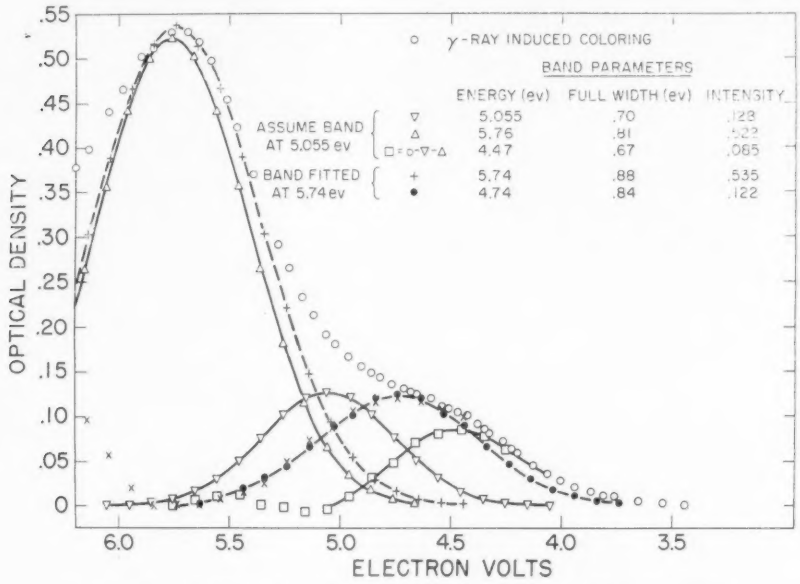


FIG. 3. The coloring induced by a gamma-ray irradiation of $2{\cdot}6 \times 10^8$ r. Two procedures have been used to separate the data into separate absorption bands which are assumed to be Gaussian shaped. The analysis resulting in three separate bands, with the suggestion of an additional one near 6 eV, agrees best with other measurements.

curve with the largest possible α_m, to the observed data, and then to fit a Gaussian to the curve obtained by plotting the difference between the first Gaussian and the data. It is seen that the data can be reasonably separated into absorption bands with maxima at 5·74 and 4·74 eV. This first resolution which resulted in three bands is slightly better than the second containing only two bands.

From Fig. 1 it is apparent that the absorption spectrum induced by a reactor irradiation at 170°C is only slightly different from the gamma-ray induced spectrum. Thus it is not surprising that when it is analyzed by the two procedures described above both a two band and a three band resolution are obtained with the peak energies and band widths close to those resulting from gamma-ray coloring.

The absorption spectra obtained by irradiating in the reactor at different temperatures show that the relative intensity of the 5·05 eV band increases regularly as the irradiation temperature is lowered. There is no evidence to support a contention that the 5·05 eV band does not contribute to the spectrum resulting from a 170°C irradiation which is quite similar to the spectrum obtained by irradiating with gamma-rays at 30°C. Because of this similarity it is reasonable to expect that if the spectrum obtained by reactor irradiation contains a component from the 5·05 eV band the gamma-ray induced spectrum should also. It has been suggested[5] that this band is formed only by reactor irradiations. If this proves to be correct it supports the analysis of the gamma-ray induced spectrum which includes only the 4·7 and 5·7 eV bands, a conclusion that makes it difficult to include a band at 4·5 eV for which there is good evidence. Thus, from our present data, we cannot rule out the possibility that the 5·05 eV band is formed during gamma-ray irradiation.

There is one phenomenon expected when materials are subjected to reactor irradiation that we are particularly interested in observing. This is the formation of imperfections, such as vacancies, and their subsequent conversion to color centers by electron trapping. It is expected that when a

vacancy is formed by the damaging process its conversion to a color center will depend on the ionization electron concentration and unless this is very high there may be an appreciable time lag between these two events. One may expect to observe this process by first irradiating a substance in a reactor and then later with purely ionizing radiation. The coloring produced by the latter irradiation should be proportional to the number of vacancies produced by the reactor irradiation and to the amount of ionization radiation providing that none of the processes involved are close to saturation. Also, the coloring produced by both reactor and ionizing radiation may be partially or entirely unstable and this will complicate the situation. A detailed analysis of the coloring to be expected in this and similar situations has been completed and will be published elsewhere. As might be expected the coloring kinetics are quite complicated. Similar phenomena are expected for other imperfections which can be formed by reactor radiations.

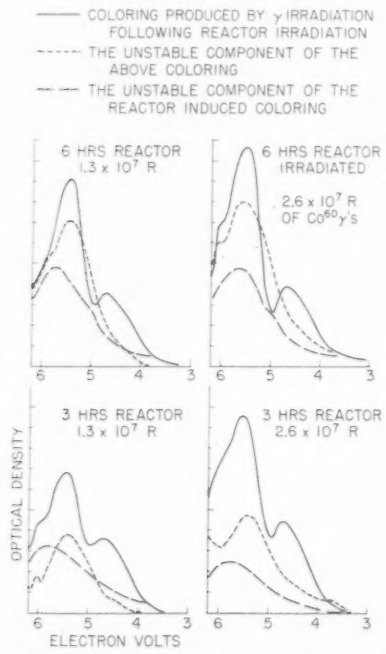

COLORING PRODUCED BY γ IRRADIATION FOLLOWING REACTOR IRRADIATION
THE UNSTABLE COMPONENT OF THE ABOVE COLORING
THE UNSTABLE COMPONENT OF THE REACTOR INDUCED COLORING
6 HRS REACTOR 1.3 x 10⁷ R
6 HRS REACTOR IRRADIATED
2.6 x 10⁷ R OF Co⁶⁰ γ's
3 HRS REACTOR 1.3 x 10⁷ R
3 HRS REACTOR 2.6 x 10⁷ R
OPTICAL DENSITY
6 5 4 3
ELECTRON VOLTS

FIG. 4. The spectra shown were obtained with four different samples that were irradiated in the reactor (the reactor induced spectra are not shown as they are very similar to the one shown in Fig. 1 for a 70°C irradiation) for the times indicated. In each case the coloring that disappeared while the sample stood at ~22°C is shown. After this unstable component had decayed to a constant value the samples were exposed, as indicated, to Co^{60} gamma-rays which produced the additional coloring plotted. Part of the coloring created by the subsequent gamma-ray irradiation is unstable and this part is also shown. These curves clearly indicate that the reactor irradiation has formed "colorable" centers in the samples that are colored by subsequent exposure to ionizing radiation and that the extent of this coloring is roughly proportional to both the reactor and gamma-ray irradiations.

To determine if a reactor irradiation does enhance the coloring of Corning fused silica when it is exposed subsequently to ionization radiation we performed the following experiment. Two, of four identical specimens, were irradiated for 3 hr in the reactor, and the other two were irradiated for 6 hr under the same conditions. The absorption spectra were measured as soon as practical after irradiation and at intervals of 1, 2 or 3 days thereafter until the unstable component of the reactor coloring had decayed away. The stable spectra remaining were very similar to the one shown in Fig. 1 for 70°C irradiations. The unstable component for all four samples is shown in Fig. 4. Because this component is small, i.e. the difference between two curves that differ only slightly, it was not susceptibile to a reliable resolution into Gaussians. By properly choosing radiation conditions these changes can be made large enough to analyze and it is our intention to deal with them in a later paper. After all four samples stood long enough to insure that only negligible changes would be obtained by further anneal at this temperature, they were then subjected to the gamma-ray doses shown in Fig. 4. The gamma-ray coloring shown in this figure is that superimposed on the reactor induced coloration. It is immediately apparent that the intensity of the coloring induced by these irradiations depends on both the initial reactor irradiation and subsequent gamma-ray coloring. In fact, this coloring is roughly proportional to both the reactor and gamma ray irradiations in accord with our speculation.

The coloring added by a gamma-ray dose of $2{\cdot}6\times10^7$ r to the stable component of coloring formed by a 6 hr reactor irradiation is shown in Fig. 5. A comparison of the four curves of this type shown in Fig. 4 indicates that the gross features are the same but that the intensities of the various bands may differ considerably. Possibly one

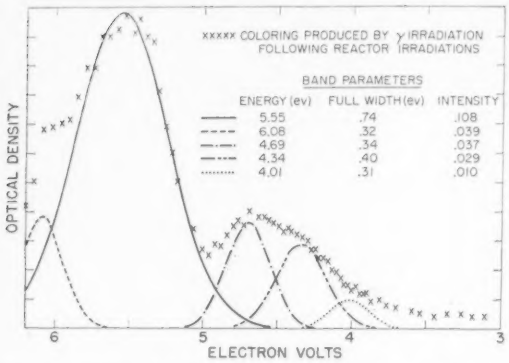


FIG. 5. Details of the solid curve shown in Fig. 4 for the sample irradiated 6 hr in the reactor and subsequently with $2{\cdot}6\times10^{7}$ r of Co^{60} gamma-rays, showing the coloring produced by the gamma-ray irradiation. Again this curve has been separated into component absorption bands assumed to be Gaussian shaped. Note that this data indicates there are two bands between 4 and 5 eV. This could be due to a poor measurement of a single band at 4·5 eV, or indicate that the two bands shown superimpose to appear as a single band, or possibly there are three separate bands in this energy region.

could use this effect to obtain information on the relative effectiveness of reactor and ionizing radiation in forming these various bands. These four curves represent four determinations of the single curve of Fig. 5, which indicates that this data is more reliable than is apparent from an examination of the one set shown. This curve was fitted with Gaussians starting with the largest peak.

It is interesting that the 5·5 eV peak, which is the most prominent band revealed by this experiment, is not apparent in the spectra of samples colored in a series of reactor or gamma-ray irradiation experiments. Thus we must consider the possibility that all of the spectra contain absorption bands that we have not detected. The curve obtained by subtracting off a 5·55 eV band requires that one consider seriously that there are two bands, and possibly even a third, between 4 and 5 eV. The presence of a band in the vicinity of 4·45 eV is shown in Figs. 2 and 3. Thus Fig. 5 could result from a poor measurement in which a single band was distorted or alternatively, only the data of Fig. 5 are sufficiently accurate to give partial resolution of the two bands. In addition, with the present data we cannot rule out the possibility that three bands may exist in this region; however, the existence of the weak band at approximately 4 eV is questionable. Note that the intensity

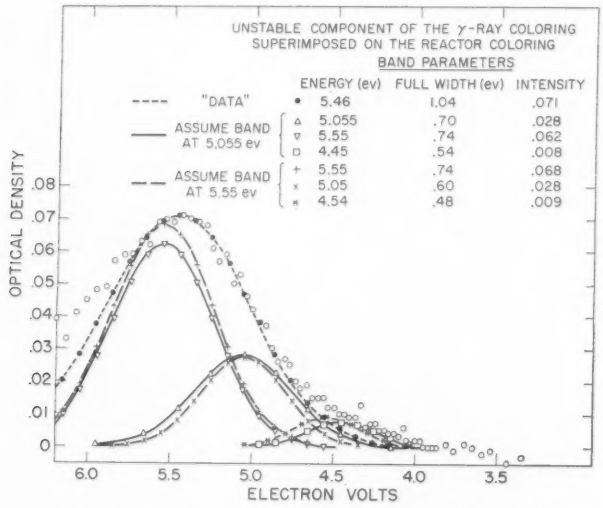


FIG. 6. That part of the coloring shown in Fig. 5 that is unstable at room temperature separated into three different sets of component bands. Either of the separations showing three bands is in reasonable agreement with other measurements.

around 4·5 eV is approximately the same in all four curves of Fig. 4. This could occur if there were 1, 2 or 3 bands between 4 and 5 eV that were due to impurities which had been irradiated to saturation, but this is not definitely established. That a band exists around 6 eV is apparent. However, the end of the operating range of the spectrophotometer is 6·18 eV and data in this range cannot be considered reliable enough to yield a significant E_0- or U-value for this peak.

Approximately one tenth of the coloring that the gamma-ray irradiation adds to the reactor irradiation is unstable at room temperatures. This unstable component is shown in Fig. 6 where three separate ways of fitting Gaussian curves to the data are shown. One can conclude that this spectrum does not require the postulation of still another band and that it is consistent with the resolutions already obtained. Since this spectrum has low intensity the parameters obtained from it are subject to large errors.

When samples of Corning fused silica are irradiated in the reactor, a weak absorption band with peak energy around 2·0 eV is formed. This band is not intense enough to estimate its full width. However, at saturation its intensity is at least 3 times "background" and it occurs in all the reactor irradiated samples.

SUMMARY

The absorption spectra obtained when Corning fused silica is colored by reactor and gamma-ray irradiations, depends on the type of radiation used and the temperature of the sample during irradiation. Assuming that the same absorption bands are present in all samples but their intensities depend on irradiation conditions, it is possible to separate, or resolve, some of the component bands from the numerous bands that are included in the observed spectra. It appears that some of the bands so resolved can be fitted with a Gaussian curve and an argument indicating that the absorption bands in fused silica could have this shape is included. Consequently, the observed spectra have been fitted to Gaussians by trial and error; a procedure—and this must be emphasized—that does not always yield a unique resolution into component bands. When samples that have been colored by reactor irradiation are exposed to gamma-rays, additional coloring is observed, and the intensity of this coloring depends on both the amount of initial reactor radiation and the subsequent gamma-ray exposure. In other words, when the samples are reactor irradiated, centers are produced that are subsequently colored when they are subjected to ionizing radiation. Some of the induced coloring, in this and in other cases, is unstable at room temperature, and curves showing the unstable components can also be analyzed by fitting Gaussian curves. Usually it is found that the unstable part can be attributed to decay of the same bands found in other spectra.

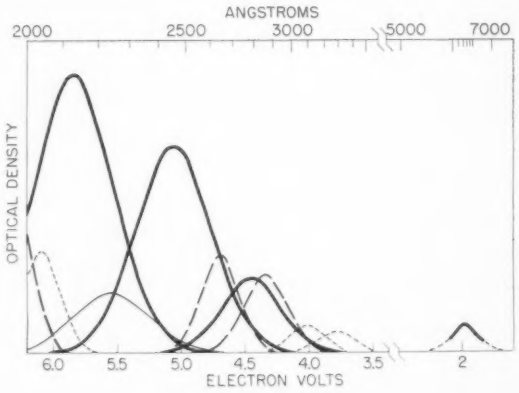


FIG. 7. A graphical summary of the absorption bands formed in Corning fused silica by reactor and gamma-ray irradiations. There is strong evidence that all of the bands shown by solid lines are present. However the band at 4·5 eV may be two unresolved bands, those at 4·4 and 4·7 eV, or actually all three bands may exist. Evidence for the band at 5·5 eV is less strong than for the other bands shown in solid lines. Evidence for the bands drawn as heavy dashed lines is greater than that for the bands drawn with narrow dashed lines. There is strong evidence for a wide band at 2 eV whose full width at half maximum is difficult to ascertain.

The various absorption bands we have found in Corning fused silica samples exposed to gamma-ray and reactor radiations are graphically summarized in Fig. 7. The solid lines represent color centers that have been separated from the observed spectra, in one way or another, with only minor reservation about the validity of the separation or their peak energy E_0 or full width U at half maximum intensity. The most serious reservation

concerns the band at 4·5 eV, shown as a single band, that might actually be the superposition of the two bands at 4·7 and 4·4 eV. However, our present data does not rule out the possibility that all three bands may exist. The evidence for the band at 5·5 eV is slightly less than for the other bands shown with heavy lines. There must be at least one weak band near 3·7 eV and another near 6·1 eV although the latter one may result from spurious measurement since it occurs at the end of the operating range of our spectrophotometer. The various dashed lines represent absorption bands that arise in the analysis given, but conclusive evidence for their existence is lacking.

Acknowledgements—Many helpful discussions with G. J. DIENES, G. H. VINEYARD and J. E. MAPES are gratefully acknowledged.

REFERENCES

1. LEVY P. W., *J. Chem. Phys.* **23**, 764 (1955).
2. COHEN A. J., *J. Chem. Phys.* **23**, 765 (1955).
3. NELSON C. M. and CRAWFORD J. H., *Bull. Amer. Phys. Soc.* **1**, 214 (1956).
4. DEXTER D. L., *Phys. Rev.* **100**, 48 (1956).
5. COHEN A. J., *Phys. Rev.* **105**, 1151 (1957).

VOL.
13
1960

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 13. pp. 296–305. Printed in Great Britain.

OPTICAL ABSORPTION IN IRRADIATED QUARTZ AND FUSED SILICA*

C. M. NELSON and J. H. CRAWFORD, Jr.

Solid State Division, Oak Ridge National Laboratory,† Oak Ridge, Tennessee

(*Received* 27 *May* 1959)

Abstract—Optical absorption studies over the range from 1850 Å to 26,000 Å have been performed on crystalline quartz and fused silica after exposure to both fast neutrons in the reactor and Co^{60} γ-rays. The prominent band near 2150 Å is produced by γ-rays and neutrons in fused silica but fast neutron exposures in excess of 5×10^{18} cm^{-2} are necessary to cause appreciable development of the corresponding band in crystalline quartz. With shorter neutron exposures and γ-rays the absorption in quartz crystals is characterized by the broad visible bands and continuous absorption which rises almost linearly toward shorter wavelength below 3000 Å. This continuous absorption probably corresponds to the long wavelength tail of one of the bands observed by MITCHELL and PAIGE in the vacuum ultraviolet region. These results do not seem to agree with their interpretation of the nature of the 2200 Å band in quartz. On annealing neutron-irradiated quartz (5×10^{18} neutrons/cm^2) at 250°C a new band at 2000 Å develops. In fused silica, besides the 2150 Å band, the shoulder at 2400 Å (resolved by appropriate optical bleaching as a band at 2570 Å) which has previously been reported, was observed with *both* γ-rays and fast neutrons. The rates of coloration of quartz and fused silica as well as studies of optical and thermal bleaching are discussed. In the fused materials a mechanism of photolytic coloration which appears to be consistent with observations provides a suitable model which is based on the rupture of Si–O bonds with the formation of free radicals as the principal product. These and other possible defects are discussed in the light of magnetic susceptibility and electron spin resonance studies.

1. INTRODUCTION

ALTHOUGH considerable effort has been spent in recent years on the study of optical absorption in irradiated silica, both crystalline and vitreous, the nature of the defects giving rise to optical absorption beyond the fundamental absorption edge is far from completely understood. This effort, however, has served to clarify the situation considerably and to highlight those more immediate aspects of the problem which must be resolved before further progress can be made. Some of the reasons for the slow progress relative to that which has been made for similar effort in the alkali halides are the much more complex crystal structure of quartz and the fact that the chemical binding is neither completely covalent nor completely ionic but lies somewhere in between. A second, no less important factor is the difficulty of obtaining specimens of high purity. The latter difficulty has been solved to an appreciable extent in the case of fused silica in that material made by a process developed by Corning shows almost none of the visible coloration which has, therefore, been attributed to impurity. In fused silica, however, there is an additional complication. In a disordered net-work structure such as this, the concept of lattice defects is no longer appropriate to describe additional radiation produced disorder and ideas of color centers based on a discrete lattice must be extensively modified before direct application is possible.

Although the problem is indeed complex, the situation is far from hopeless as the papers presented at this symposium amply attest. Moreover, attention should be called to the recently reported

* Paper presented at *Symposium on Defect Structure of Quartz and Glassy Silica*, Mellon Institute, May 22–24, 1957.

† Oak Ridge National Laboratory is operated by Union Carbide Corporation for the U.S. Atomic Energy Commission.

VOL. 13 1960

work of MITCHELL and PAIGE[1] on quartz. They extended the spectral range in the ultraviolet to the absorption edge (1460 Å) and have discovered two additional bands, the *D*-band at 1730 Å in γ- and X-irradiated crystals and the *E*-band at 1630 Å in neutron irradiated crystals. Studies of coloration vs. neutron dose, the behavior of γ-irradiated specimens, thermal and optical bleaching behavior and thermal annealing have lead them to the conclusion that the *E*-band is produced by a hole trapped at an interstitial oxygen atom and the *C*-band (2200 Å in their notation) is caused by an electron trapped on an oxygen vacancy. These conclusions are based on rather indirect evidence. However, their model is the first that has been proposed which includes all of the optical data with the exception of the so-called *D*-band, and is quite valuable in that it presents a point of departure from which to orient further thinking. Their work also points up the importance of covering the entire spectral range. It may be said without fear of contradiction that a realistic model of radiation produced defects in quartz and fused silica must await a careful study of the vacuum ultraviolet region.

In addition to absorption measurements in the vacuum ultraviolet range certain correlary studies are also vital to the correct interpretation of color centers in the silica system. Perhaps the greatest aid may be expected to come from magnetic studies of the defects. Electron spin resonance measurements have the advantage of being interpretable in microscopic terms and should therefore yield much information about the nature of individual defects. The identification of the visible absorption bands (*A*-bands) in irradiated quartz with ionized aluminum impurity was a direct result of correlating ESR results with optical absorption spectra.[2–4] Although a similar correlation between the ultraviolet absorption bands and ESR lines is yet to be made, studies of this type are showing considerable progress.[5] Additional correlary information from magnetic susceptibility measurements[6] and studies of thermoluminescence should also serve a valuable function in the construction of a model of color centers.[7,8]

The work described here was initiated in an attempt to learn more about the chemical and electronic effects that must accompany the loss of structure in quartz after prolonged neutron irradiation. To this end measurements of optical absorption spectra, paramagnetic resonance and magnetic susceptibility on various forms of silica as a function of neutron and γ-ray exposure were undertaken at Oak Ridge. Although the original aim is yet to be reached, much has been learned about the defects produced by energetic radiation. Included here is a summary of the rate of coloration with both γ-rays and neutrons, the optical and thermal bleaching behavior and the effect of previous bombardment on the susceptibility of fused silica to γ-ray coloration.

The neutron exposures were performed in a partial fuel element in the lattice of the Low Intensity Test Reactor (LITR) and a Co^{60} γ-ray source of 4×10^5 Mr/hr intensity was used for the γ-ray exposures. Optical absorption measurements were made with a Cary Model 14 spectrophotometer. This instrument records the spectra directly in terms of optical density and its short wave length limit is 1850 Å.

2. EFFECTS OF NEUTRON EXPOSURE

The optical absorption spectra of three types of silica after a small dose of reactor radiation ($\sim 10^{17}$ fast neutrons/cm^2) are shown in Fig. 1 in which extinction coefficient (per mm) *E* is plotted against wavelength over the range from 1850 to 4000 Å. Corning silica is the most resistant to

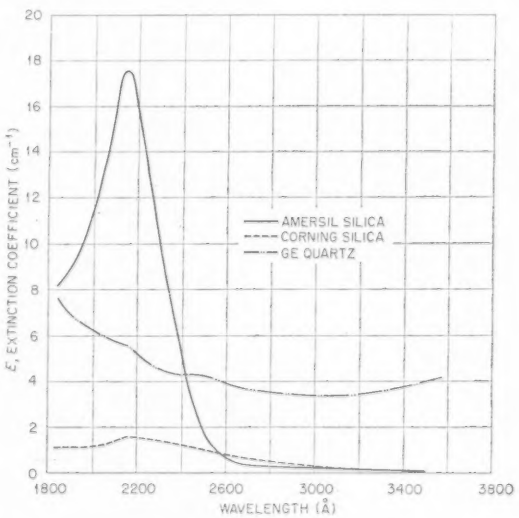


FIG. 1. Absorption spectra of irradiated SiO_2 samples (1×10^{17} nvt).

coloration and the Amersil optical grade No. 1 is the least. While both varieties of fused silica show the familiar 2150 Å absorption peak of *C*-band there is a factor of ten difference in absorption at this wavelength. Intermediate between these is quartz, in this case General Electric Ltd. synthetic quartz, though natural quartz gives a similar absorption in this wavelength region. In addition to the *A*-bands located beyond 4000 Å, there is evidence of a weak band at $\sim$ 2500 Å and the 2200 Å *C*-band is faintly visible. These bands are superimposed on a continuous background absorption which rises rapidly toward shorter wavelengths. This may be the tail of one of the vacuum ultraviolet bands observed by MITCHELL and PAIGE, i.e. the *D*- or *E*-band. The 2500 Å band is not usually observed in quartz specimens and it is therefore concluded that it is an impurity band. The development of the so-called *C*-band for such short neutron exposures is quite variable from specimen to specimen, usually being more intense in natural crystals. In no case, however, was the so-called *C*-band anywhere near as intense as the *A*-band complex.

After prolonged neutron bombardment ($> 10^{20}$ fast neutrons cm^{-2}) the introduced damage outweighs previous specimen history. This is demonstrated by the spectra of the various forms of silica. These are shown in Fig. 2 for specimens of the same type as included in Fig. 1 after an exposure of $2{\cdot}9\times10^{20}$ fast neutrons cm^{-2}. The spectra of the three specimens are quite similar, the only difference being a small variation in the magnitude of the extinction coefficient. The predominant feature is the 2150 Å peak (*C*-band) which now falls at the same position for the quartz as well as the fused silica. Faintly evident is the band near 2500 Å, the B_2-band in the nomenclature of MITCHELL and PAIGE[1]. Both of these bands are superimposed on a background absorption which is presumably the tail of the *E*-band (1620 Å) observed by MITCHELL and PAIGE. It is interesting to note that the *A*-band absorption in the quartz sample has been completely bleached during the heavy neutron exposure in agreement with previous observations.[1,9] The conversion of the quartz spectrum to one almost identical to the glasses is to be expected on the basis of structure studies since the X-ray diffraction patterns of heavily irradiated quartz ($\sim 10^{20}$ fast neutrons cm^{-2}) indicate that all long range order has been destroyed and the density of both fused and crystalline material approach the same saturation value.[10]

The intensity of coloration as a function of neutron dose is shown in Fig. 3. The absorption coefficient of 2150 Å is plotted against integrated fast neutron flux. All specimens tend to saturate near the same absorption. The maxima in the curves for quartz are to be noted. These may be related to the loss of crystallinity which occurs over the same exposure range. The rate of approach to saturation does not obey a simple first order law even in the case of fused silica. Instead the curves show a complex behavior which may

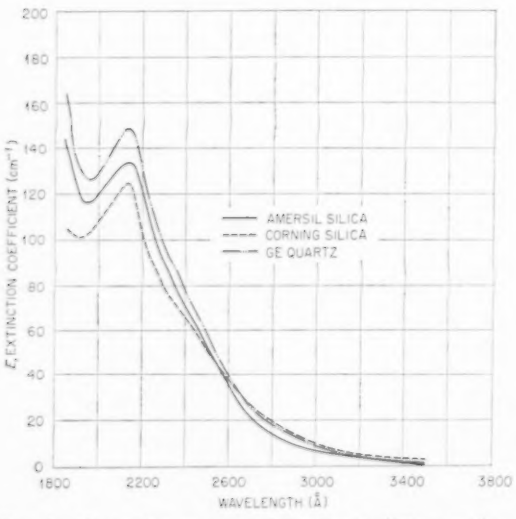


FIG. 2. Absorption spectra of irradiated SiO_2 samples ($2{\cdot}9\times10^{20}$ nvt).

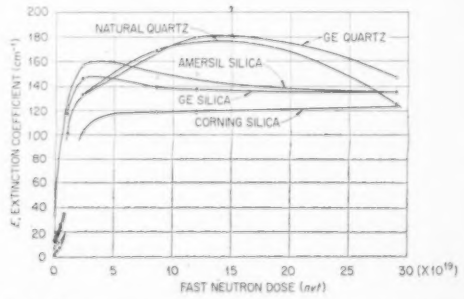


FIG. 3. Fast neutron bombardment of SiO_2 samples.

indicate that some of the initial coloration may be produced by centers which become unstable as the bombardment proceeds. In view of LEVY's observations on the complex nature of the *C*-band,[11] such an interpretation does not seem unreasonable for the fused specimens and in quartz it is well known that the *A*-bands, which contribute appreciably to the 2150 Å absorption, are bleached by prolonged neutron exposure. A similar result has been obtained by MITCHELL and PAIGE in the case of quartz specimens. LEVY has treated the shape of coloration vs. exposure curves in some detail.[12]

VOL. 13 1960

3. γ-RAY COLORATION

The absorption spectrum of quartz colored to near saturation by Co^{60} γ-rays ($\sim 3\times10^8$ r) is shown in Fig. 4. The solid curve relates to a

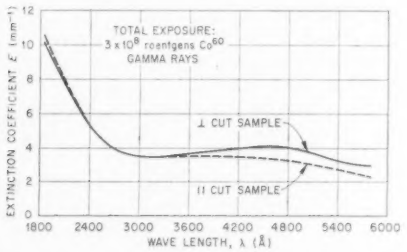


FIG. 4. Absorption spectra of GE synthetic quartz after gamma irradiation.

General Electric Ltd. crystal in which the *C*-axis is perpendicular to the face of the plate while the dashed curve was obtained on a plate cut parallel to the *C*-axis. These specimens are designated as $C_\perp$ and $C_\parallel$, respectively. It should be noted that this notation is the inverse of that employed by MITCHELL and PAIGE[4] who used the electric vector of the incident beam as the reference direction. Although these samples were obtained from the same batch, the impurity content may have been somewhat different and we cannot say with certainty that the parallel orientation absorbs light more strongly at short wavelengths than the perpendicular orientation. In two other measurements on pairs of specimens cut from both natural and synthetic crystals it was found that the absorption in the parallel orientation was considerably less than for the perpendicular orientation below 2500 Å. It should be mentioned that MITCHELL and PAIGE, who used highly polarized light, found no significant difference in the absorption in this range for the two crystal orientations. The most striking feature of the spectrum in Fig. 4 is the absence of any evidence of the so-called *C*-band at 2200 Å. At much lower exposures there is a hint of such a band but its intensity is not at all comparable to that of the *A*-band complex. In natural quartz this band is somewhat better developed. However, even here an examination reveals that the area under the peak, after subtraction of the background, is only of the order of one per cent of the area under the *A*-band system. Therefore, if the *A*-bands and the *C*-band are coupled, as MITCHELL and PAIGE[1] suggest, the oscillator strength of the 2200 Å band is only about one per cent of the value for the combined *A*-bands. Consequently, we question the interpretation given by MITCHELL and PAIGE of the nature of the 2200 Å band based on an *A*-band–*C*-band interrelationship, namely that the *C*-band in quartz is produced by electrons trapped at oxygen vacancies, these electrons originating at substitutional aluminum impurity. It should be pointed out that an examination of their data does not indicate conclusively that a band of appreciable amplitude exists at this position after γ-ray exposure. The limit of the range of the two instruments used overlaps in this range. Therefore, the large apparent maximum they observe may be distorted by instrumental limitations. The data shown in Fig. 4 were obtained on a Cary Model 14 spectrophotometer which covers the range from 1850 Å to 4000 Å continuously. It should be mentioned in passing that this spectrum is quite similar to most of those obtained on synthetic quartz after short reactor exposures ($< 5\times10^{17}$ cm^{-2}).

The spectrum resulting from γ-ray exposure in Corning silica is shown in Fig. 5. The solid curve without points was obtained after exposure to 5×10^8 r of Co^{60} γ-rays. This corresponds to near saturation coloration (see below). Also shown for comparison is the spectrum obtained after exposure to 1×10^{20} fast neutrons cm^{-2} normalized to coincide with the γ-ray curve at the 2150 Å maximum. This comparison shows that neutron exposure results in (i) a greater development of the far ultraviolet band, (ii) a greater width of the 2150 Å band, and (iii) considerably less development of the 2550 Å band. These differences bear

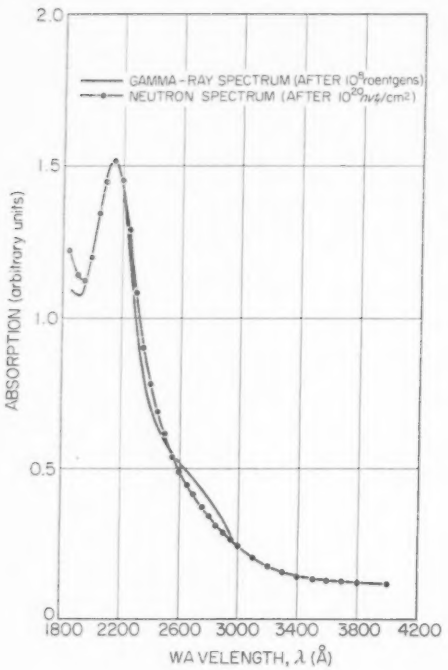


FIG. 5. Absorption spectra of Corning silica after neutron and gamma-ray irradiations.

out LEVY's conclusion that absorption in the region from 2000 to 3000 Å is made up of a number of overlapping bands.

The rate of coloration of both quartz and fused silica during γ-ray exposure plotted in Fig. 6 shows some interesting features. Coloration vs. exposure for an Amersil optical grade No. 1 specimen measured at 2150 Å shows a simple initial build-up to an early apparent saturation near 10^7 r. Continued exposure, however, causes a slow linear increase, and the slope of this portion of the curve is almost the same as that for the Corning specimen. Measurements to exposures $> 10^9$ r reveal that this slow increase persists in the two types of silica. Consequently, we assume that this is due to creation of defects by the γ-rays, presumably displacement of atoms by the energetic Compton electrons, a situation observed in germanium.[13] The Varley mechanism[14] in which bond rupture and displacement follow multiple ionization of the electronegative constituent is also a distinct possibility.

The absorption of quartz at 1850 Å as a function of exposure is shown in the remaining curve of Fig. 6. Once again complex initial behavior is noted. The reason for the early sigmoid curvature is unknown. It looks as though a certain induction

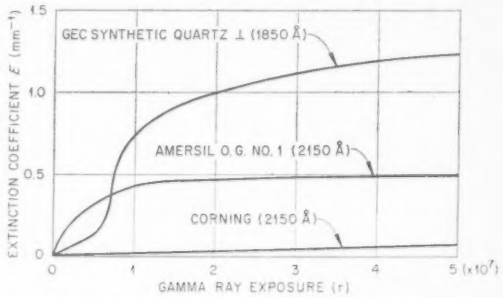


FIG. 6. Build-up of coloration in SiO_2 during gamma-ray exposure.

period is required before the rate of coloration becomes appreciable. One possible explanation is that certain impurity traps with a large capture cross section need to be filled before the normal radiation coloration can proceed.

4. OPTICAL BLEACHING

Much of the coloration in irradiated silicas can be bleached by light falling in the range from 2000 to 3000 Å. In γ-irradiated *quartz* almost all of the absorption above 1850 Å can be bleached with a 500 W mercury vapor lamp whose peak intensity falls in the 2537 Å range and which also emits appreciable intensity near 2000 Å. After successive periods of bleaching little if any spectral selectivity can be observed and bleaching is apparently uniform over the spectral range in question (1850–4000 Å). After neutron exposure, bleaching of quartz is more difficult, less coloration being removed for a given bleaching period the longer the exposure. This suggests that the defects produced by bombardment are either more difficultly bleached or that these centers have a high cross section for recapturing electrons. After prolonged neutron exposure ($\sim 10^{20}$ cm^{-2}) the bleaching characteristics of quartz specimens are the same as for heavily irradiated fused silica as expected.

In contrast to quartz γ-irradiated fused silica is somewhat more resistant to optical decoloration. However, the absorption at 2150 Å can be reduced

VOL. 13 1960

a factor of 3–4 by prolonged bleaching with ultraviolet light. A series of bleaching experiments were carried out at liquid nitrogen temperature and it was found that the bleaching efficiency was considerably reduced at the low temperature. The decoloration produced by an exposure to ultraviolet light was smaller by about a factor of 5 at 77°K as compared to room temperature bleaching for an equal period and intensity. This behavior suggests that the excited state of the absorbing center is a discrete state and does not lie in the ionization continuum. The alternative would require a thermally activated destruction of the center after its excitation.

Irradiation with ultraviolet at 77°K also causes a pronounced change in the spectra. Fig. 7 shows the change in absorption produced by prolonged exposure to the 500 W mercury vapor lamp. There is a marked increase in the absorption to the long wavelength side of the 2150 Å peak, whereas a decrease on the short wavelength side is observed. The absorption increase suggests that a subsidiary band near 2200 Å is either developed or enhanced by a redistribution of electrons over various possible defects.

This spectral shift observed during bleaching at 77°K has been shown to depend on the spectral distribution of bleaching light in the case of room temperature studies. For example use of a 2 W germicidal lamp (very little intensity below 2537 Å) causes a marked decrease in absorption at 2150 Å while affecting the band near 2550 Å only slightly. Consequently, such a bleaching exposure makes more distinct the longer wavelength subsidiary band as shown in Fig. 8. If, now, after a short bleaching exposure to the germicidal lamp,

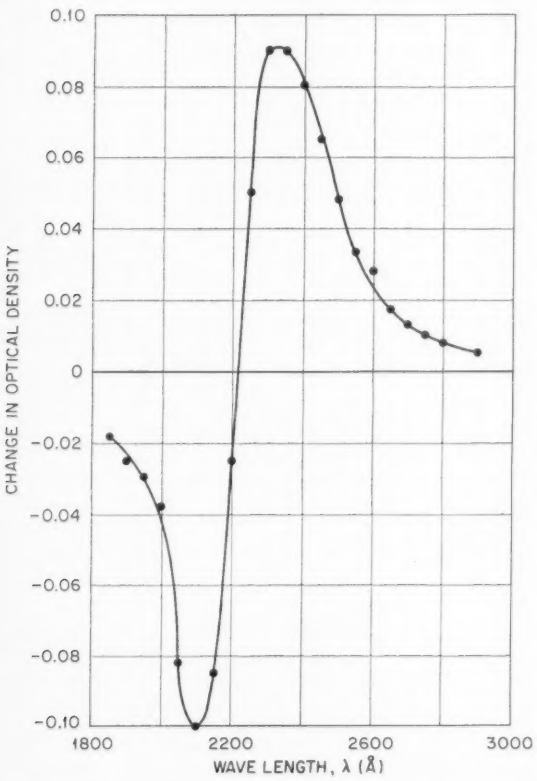


FIG. 7. Optical bleaching of gamma-irradiated Corning silica after $1{\cdot}5\times10^{9}$ r Co^{60} gamma-rays. Exposed to ultraviolet light at liquid nitrogen temperature for 2 hr.

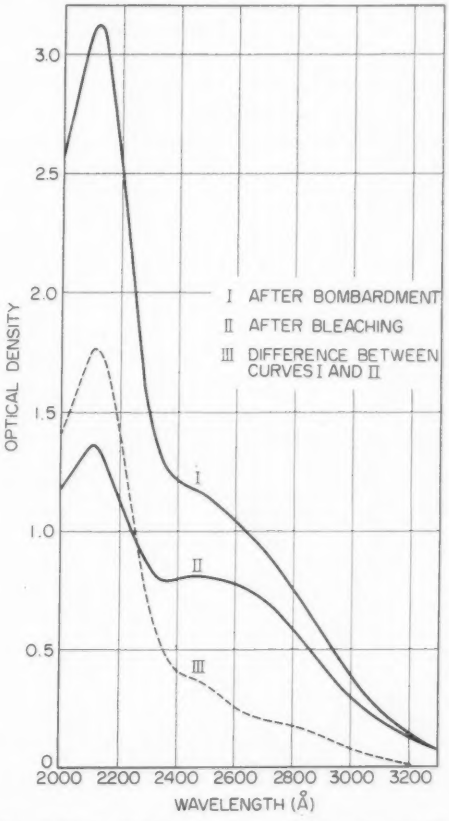


FIG. 8. Optical bleaching of neutron-bombarded Corning silica.

a neutron bombarded Corning silica specimen is further exposed to the 500 W lamp the absorption to the long wavelength side of the 2150 Å maximum is increased even though the absorption at 2150 Å continues to decrease. This behavior is quite similar to the observation after liquid nitrogen bleaching. Aside from the greater intensity of the 500 W lamp, the major difference between it and the germicidal lamp is a greater relative intensity between 2000 Å and the 2537 Å line. Consequently, it appears that electrons or holes may be excited from a band to the short wavelength side of the maximum into a defect leading to absorption between 2150 Å and 2600 Å, a behavior that is consistent with LEVY's picture of the *C*-band complex.[11]

5. THERMAL ANNEALING OF COLORATION

In addition to optical bleaching, thermal bleaching studies of the coloration in irradiated SiO_2 specimens have also been initiated. For γ-irradiated quartz and silica specimens essentially all absorption can be removed by heating to 350° to 400°C. After appreciable neutron exposure ($> 10^{18}$) the thermal decoloration process becomes much more complex and higher temperatures are required for complete decoloration. Fig. 9 shows the effects of thermal annealing on a synthetic quartz specimen after exposure to 3×10^8 r of Co^{60} γ rays. Optical density is plotted against the anneal temperature. The specimen was held at each anneal temperature for a period of one hour. The upper curve refers to the absorption at 2500 Å and the lower curve to absorption at 4000 Å. Very sharp decoloration occurs over the range from 200° to 300°C suggesting that a singly activated process is responsible for the bleaching. This behavior is consistent with the results of MITCHELL and PAIGE[1] who find an activation energy of 0·75 eV for the thermal bleaching process from an analysis of isothermal bleaching curves. It should be further pointed out that the 2500 Å curve and the 4000 Å curve are not directly superimposable, the former showing continued curvature at high temperatures. This suggests that the annealing of the *A*-bands proceeds at a different rate than that of the ultraviolet absorption bands.

Thermal decoloration of neutron irradiated quartz is more difficult and more complex than that for γ-irradiated specimens. Heavily irradiated specimens behave in the same way as neutron bombarded fused silica (see below). For small and moderate bombardments a complex change in the absorption spectrum is produced by heating. Fig. 10 shows the absorption spectrum of a synthetic quartz specimen exposed to 5×10^{17} fast neutrons cm^{-2} after successive heat treatments. After exposure there is evidence for a slight development of the 2200 Å *C*-band and indications of a weak

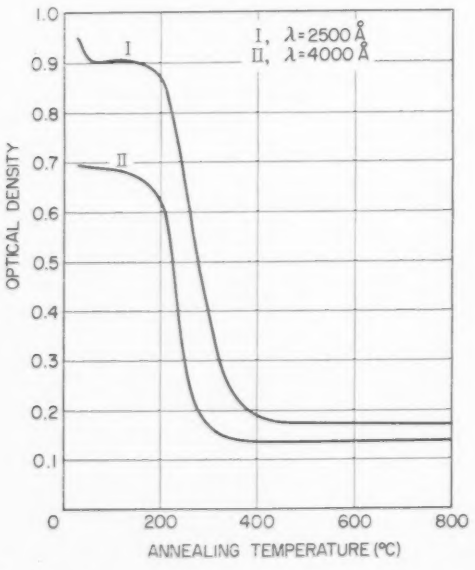


FIG. 9. Annealing of coloration in gamma-irradiated quartz. Optical density vs. annealing temperatures (~1 hr at each temperature).

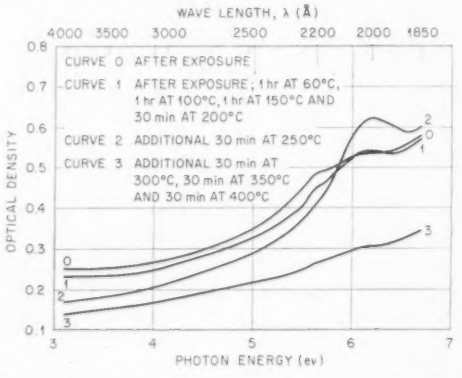


FIG. 10. Optical density vs. photon energy after neutron exposure ($nvt_f \simeq 5\times10^{17}$) and anneals.

VOL. 13 1960

absorption at 2000 Å. Both of these are superimposed on a strong background absorption which rises continuously toward short wavelength. On heating successively at 60°C, 100°C and 150°C for one hour and at 200°C for one-half hour, the background absorption and absorption in the 2200 Å region is noticeably diminished while there is a small enhancement in the 2000 Å band. After an added anneal at 250°C, the band at 2000 Å becomes much enhanced. Annealing at successively higher temperatures removes all structure leaving only a greatly diminished background absorption. The nature of the 2000 Å band is not known. It may be the same as the one at this position postulated by LEVY in fused silica.[11] It is significant, however, that after extensive bombardment disordering of the quartz structure (heavy doses), enhancement of absorption in this range is no longer evident. Therefore, it must be concluded that if this band is indeed produced by thermal treatment in heavily irradiated specimens it is obscured by absorption at other centers and, hence, its production by irradiation is not nearly as efficient as the latter.

The thermal bleaching of a γ-irradiated Corning specimen at 2150 Å is compared with that for a neutron bombarded one in Fig. 11. After γ-irradiation rapid bleaching in the range from 200° to 350°C is observed. However, the form of the curve suggests that two processes rather than a single one are operating. After neutron bombardment two major annealing stages are still evident though the thermal decoloration is smeared out over a much wider temperature range. This seems to indicate that the activation energy for emptying the absorption centers is spread out into a distribution after neutron bombardment. Such a situation is a common one for recovery of property changes after neutron bombardment. On the basis of these results it is impossible to draw any conclusions about the nature of the two processes postulated above. These may arise from different bleaching rates of bands which overlap in this region. Alternatively, there may be more than two processes which add to give the observed behavior. After heavy neutron exposures the annealing is spread over an even greater temperature range and the behavior of quartz and glass become quite similar.

The effect of previous neutron exposure on γ-ray colorability has also been studied. Extensive disruption of the Si–O network would be expected to cause an appreciable increase in both the rate of growth of absorption and its saturation value. Fig. 12 shows this effect for Corning silica which has received 5×10^{17} fast neutrons cm^{-2} followed by a thermal anneal at 700°C. A marked increase in coloration rate and saturation absorption at 2150 Å is observed. For comparison the curve for a previously γ-irradiated specimen after the same heat

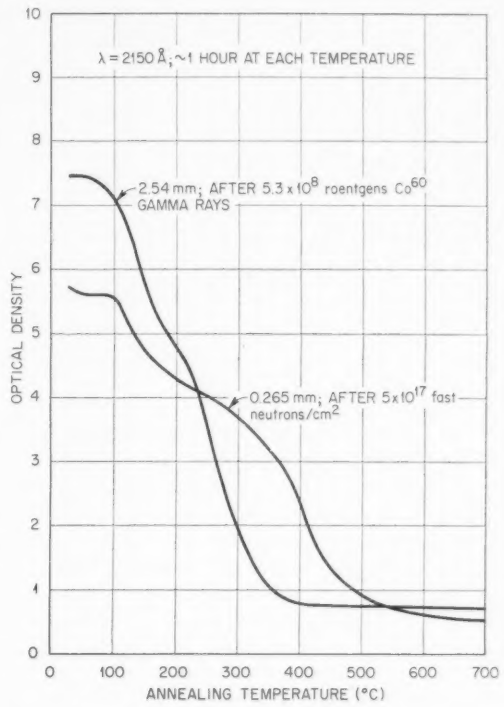


FIG. 11. Thermal bleaching of Corning silica; optical density vs. annealing temperature.

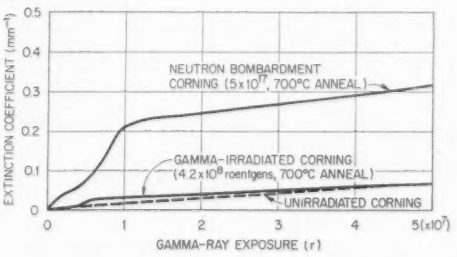


FIG. 12. Build-up of coloration in previously irradiated SiO_2.

treatment is shown. Little if any enhancement in colorability of this specimen was observed. We must conclude from this experiment that the neutron produced defects in the network are not completely annealed by heating for 1 hr at 700°C. Moreover, occupation of a given defect by an electron or a hole is necessary for optical absorption.

In contrast to the fused silica the coloration vs. dose curve for the quartz specimen shows an appreciable decrease after the same pretreatment. The reason for this behavior is quite obscure. Offhand, a similar result to that obtained in the glass would be expected.

Recoloration experiments on both crystalline and fused specimens after optical bleaching of neutron irradiated specimens have also been conducted. Once again, as would be expected short γ-ray exposures rapidly recolor the specimens to near the absorption prior to the optical bleach, presumably by refilling the existing defects. These experiments also show some interesting spectral changes similar to those discussed above. However, the results are not yet complete enough to permit a detailed discussion.

6. SUMMARY

In summary the following observations can be made. Exposure of crystalline quartz to energetic photons (Co^{60} γ-rays in the present case) introduces a diffuse system of bands in the visible range and an almost structureless absorption which rises gradually toward shorter wavelengths. It has been established that the visible bands are associated with substitutional aluminum impurity but how much of the absorption extending into the ultraviolet range is impurity connected is not known. There is some structure in the 2000–4000 Å region in the form of weak diffuse bands which is variable from one crystal to another and, therefore, must be associated with impurities. This latter type of absorption is observed primarily in natural crystals and is either very weak or absent in synthetic quartz. In marked contrast to quartz, γ-irradiated Corning fused silica—in which by virtue of its high purity the visible impurity absorption is absent—exhibits a definite absorption system with a maximum located at 2150 Å, evidence for a subsidiary band near 2500 Å, and rapidly rising absorption below 2000 Å, indicating additional structure in the far ultraviolet. The γ-ray induced absorption in both Corning silica and crystalline quartz can be largely bleached by an unfiltered mercury vapor lamp or by heating to temperatures in excess of 350°C.

On prolonged fast neutron exposure ($> 10^{20}$) the absorption spectrum of quartz is transformed into one approaching that of Corning silica. The latter material after a similar neutron exposure exhibits a spectrum which, except for a much enhanced intensity, is quite similar to that observed after γ-ray exposure. Hence it appears that the absorption system centered at 2150 Å is characteristic of the damaged glassy structure as are the ultraviolet bands reported by MITCHELL and PAIGE[1] after neutron exposure. The neutron irradiated specimens are much more difficultly bleached with either heat or ultraviolet light than those exposed to γ-rays.

A close examination of the 2150 Å system after various manipulations indicates its complex character. For example bleaching light with a large component of intensity between 2537 Å and 2000 Å causes a shift of the system to longer wavelengths, i.e. greater bleaching at short wavelengths than at longer wavelengths. This behavior is consistent with LEVY's contention that the *C*-band is a complex structure composed of several independent bands.

In view of the lack of correlation between the band at 2200 Å and the visible bands in crystalline quartz, we question the interpretation given by MITCHELL and PAIGE for absorption near 2200 Å after X- or γ-ray exposure and, therefore, the defect model of the *C*-band after neutron exposure. Since appreciable neutron exposure is necessary for a perceptibile development of this band, it seems more likely that it is associated with vitreous regions in the crystal produced by the bombardment or at least with defects not produced in high yield by γ-rays in crystalline quartz.

In high purity fused silica the origin of the γ-ray colorability in the ultraviolet region and its wide variability in different types of comparable material is puzzling. In order to explain these factors we have postulated that the coloration is associated with ruptured covalent bonds in the Si–O network.[15] Since the lifetime of such free radicals is expected to be quite short in the absence of constraint, it was further postulated that they were stable only in strained regions where, after

stress relaxation, the internuclear separation was too great for easy bond reformation. Consequently, in this picture the 2150 Å complex would be associated with a partial valence bond localized at a silicon atom or at an oxygen atom.

REFERENCES

1. MITCHELL E. W. J. and PAIGE E. G. S., *Phil. Mag.* **1**, 1085 (1956).
2. GRIFFITHS J. H. E., OWENS J. and WARD I. M., *Rep. Bristol Conf. on Defects in Crystalline Solids* p. 81. Physical Society, London (1955).
3. O'BRIEN M. C. M. and PRYCE M. H. L., *Rep. Bristol Conf. on Defects in Crystalline Solids* p. 88. Physical Society, London (1955).
4. MITCHELL E. W. J. and PAIGE E. G. S., *Phil. Mag.* **46**, 1353 (1955).
5. WEEKS R. A., *J. Appl. Phys.* **27**, 1376 (1956).
6. STEVENS D. K., *Phys. Rev.* **98**, 1541 (1955).
7. LAUTOUT M., *J. Chim. Phys.* **52**, 259 (1955).
8. ARNOLD G. W., following paper.
9. DITCHBURN R. W., MITCHELL E. W. J., PAIGE E. F. S., CUSTERS J. F., DYER H. B. and CLARK C. D., *Rep. Bristol Conf. on Defects in Crystalline Solids* p. 92. Physical Society, London (1955).
10. WITTELS M. and SHERRILL F. A., *Phys. Rev.* **93**, 1117 (1954).
11. LEVY P. W., *Color Center Symposium, Argonne National Laboratory*, November 1956.
12. LEVY P. W., *Bull. Amer. Phys. Soc.* **1**, 136 (1956).
13. CLELAND J. W. and CRAWFORD J. H., *Phys. Rev.* **102**, 722 (1956).
14. VARLEY J. H. O., *J. Nucl. Energy*, **1**, 130 (1954).
15. NELSON C. M. and CRAWFORD J. H., *Bull. Amer. Phys. Soc.* **1**, 214 (1956).

VOL. 13 1960

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 13. pp. 306–320. Printed in Great Britain.

DEFECTS IN NATURAL AND SYNTHETIC QUARTZ*

G. W. ARNOLD, Jr.

United States Naval Research Laboratory, Washington, D.C.

(*Received* 27 *May* 1959)

Abstract—Thermoluminescence measurements of quartz (both irradiated and naturally colored) have been made from room temperature to 375°C. Various peaks of the light intensity vs. temperature curves have been identified with absorption maxima in the 190–1000 mμ spectral region. The method is extremely sensitive to radiation dose; well-defined peaks being observed for very short X-irradiation times (~ 1 sec.). Data are presented on X-irradiated natural quartz, smoky quartz and on X-irradiated synthetic quartz with various additives.

INTRODUCTION

ONE of the effects that ionizing radiation can have on clear crystalline quartz and fused silica is the alteration of their optical absorption spectra in the visible and ultraviolet regions. This occurs when electrons or holes do not recombine but are trapped at lattice sites in which they may remain for long periods of time. Such spectral changes that result are often in the form of characteristic absorption maxima at well defined energies and are attributed to "color centers".

The transformation of the original unirradiated sample of quartz or silica into a state in which light quanta are absorbed in spectral regions which previously contained no absorption maxima results in an increase in lattice energy.

The release of this energy in quartz and silica can be brought about by irradiation with light which can be absorbed by certain of the centers produced by the ionizing radiation or by heating the sample. Both processes can result in returning the specimen to its original state insofar as its optical absorption spectra is concerned. (Effects due to displacing irradiation will not be discussed here.) This return to its original state requires the recombination of electrons and holes and this process may give rise to emitted quanta. When the recombination is caused by heating, the emitted quanta are termed products of "thermoluminescence". A plot of the intensity of the emitted radiation as the temperature of the irradiated sample is varied may show well defined maxima at various temperatures. Such plots are termed "glow curves" and their production and interpretation for certain phosphor systems has been dealt with in a thorough manner by RANDALL and WILKINS[1,2,3] and GARLICK and WILKINS[4].

The quantitative interpretation of glow curves for materials in which there are several trapping levels is difficult because of the constant state of flux of electrons and holes and their possibly complex interplay with trapping centers. One can, however, use the information obtained in conjunction with optical absorption data and establish a correspondence between certain of the thermoluminescence maxima and absorption centers. The advantage obtained in working with thermoluminescence curves lies in the extreme sensitivity of the phenomenon to radiation doses. Experimentally, this means irradiation times in terms of seconds or minutes instead of the hours or days often necessary in optical absorption studies of irradiated quartz and silica.

The thermoluminescence data presented in this paper were obtained and interpreted against a background of considerable investigation of the optical absorption spectra of natural and synthetic quartz. Some familiarity with the past history of the work on natural and artificial coloration of quartz and silica is therefore considered relevant and in the following paragraphs a brief, but by no means complete account is presented.

* Paper presented at *Symposium on Defect Structure of Quartz and Glassy Silica*, Mellon Institute, Pittsburgh, Pa., May 22–24, 1957.

VOL. 13 1960

BACKGROUND

(a) *Color centers*

At the present date, most of the color center work in quartz is tentatively based on the model proposed by MITCHELL and PAIGE[5]. This is chiefly due to the fact that it is the only model thus far proposed which deals with all of the observed color center maxima. It is, however, of at least historical interest to note the work leading up to the present-day state of affairs. The material in the following paragraphs of this section are in, more or less, chronological order. The facts which the present writer has chosen to extract from each referenced paper represents what he feels to be of most interest and does not necessarily represent a summary of the paper. The author has not attempted to criticize any of the referenced studies in the light of our present knowledge. DIENES (this symposium) has pointed out the large gaps which exist in our picture of silica and such criticism would be premature. The observed absorption maxima in quartz and silica are entered in Table 1.

The early work on color in quartz was reviewed by HOLDEN[6] in 1925. These investigations were chiefly attempts to correlate color in smoky quartz, amethyst, citrine and rose quartz with specific impurities acting as pigmenting substances. A fairly complete bibliography is given by HOLDEN, and some representative work is referenced in the present paper.[7–26]

HOLDEN[6] measured the transmission of light through smoky and amethyst quartz. He found absorption near 600 mμ in smoky quartz and in clear quartz colored by γ-irradiation. In amethyst he found a maximum at 530–540 mμ. His measurements were made from 450 mμ to 700 mμ.

MOHLER[27] also made absorption measurements on smoky quartz and amethyst. She found a broad maximum centered at about 425 mμ in smoky quartz. Bands at 546 mμ and 262 mμ were noted in amethyst. Further work by MOHLER[28,29] on irradiated clear quartz and natural smoky quartz produced evidence for absorption maxima at 606, 500–488, 400, 303–294 and 263 mμ in both cases. Measurements were possible from 200 mμ to 650 mμ.

FUTAGAMI[30] found bands at 390 mμ and 302 mμ in X-irradiated natural quartz. The 390 mμ absorption was strongest in those specimens which were brown in color after X-irradiation, while the 302 mμ bands was strongest in those which became violet after irradiation. It should be noted that the 302 mμ band is a recognized center in fused quartz. (See Table 1.)

CHOONG[31] made studies of γ-irradiated colorless, smoky, citrine and rose quartz. He found that X- or γ-irradiation made the clear quartz smoky, deepened the smoky color in smoky quartz and superimposed the smoky color on the natural color of citrine and rose quartz. CHOONG found no absorption maxima for the crystalline quartz but did find that the vitreous material obtained from melted crystalline specimens turned violet when irradiated and gave a band maximum at 300 mμ. He found also that ultraviolet light from a hydrogen lamp could be used to bleach the artificial smoky color. CHOONG found it possible to color smoky crystalline quartz with ultraviolet light, the exposure time (1 kW H_2 light) being of the order of 30 hr for just perceptible coloration. No coloration was noted for vitreous smoky quartz or for the crystalline and vitreous varieties of the other specimens.

VEDENEEVA[32] found an absorption maximum in amethyst at 540 mμ and noted the similarity with the absorption in glass which has been colored violet by irradiation. Further investigations were made by VEDENEEVA[33] with polarized light on smoky quartz. She found the dichroism to be greatest at 300 mμ and at 400 mμ.

SMAKULA[34] found the visible absorption peak in X-irradiated natural quartz to be anisotropic. He found the position to vary between 400 mμ and 435 mμ, depending on the crystal orientation. He believed a second band to exist below 205 mμ in both crystalline and fused quartz. In addition, fused quartz (X-irradiated) showed peaks at 540 mμ and 305 mμ.

FORMAN[35] colored sections of natural quartz cut perpendicular to the *c*-axis with X-rays and measured the absorption from 400 mμ to 600 mμ. His data show a maximum in absorption at about 400 mμ for quartz which had been colored to saturation. One sample which was neutron irradiated had a peak in the region 450–470 mμ. He found that the rate of coloration after decolorizing by baking was increased, especially if the heat applied was just sufficient to remove the color. Longer periods of heating at higher temperatures caused the crystal to recolor at the same rate as in the original irradiated quartz.

BAPPU[36] investigated amethyst quartz and found that the absorption maximum at 550 mμ behaved like an *F*-band in the alkali halides, i.e. it sharpens and shifts to shorter wavelengths with decreasing temperature and when irradiated with *F*-band light, a band similar to an *F'*-band appears at longer wavelengths. In further work[37] with amethyst in the ultraviolet and infrared regions he found bands at 785 mμ, 835 mμ and 920 mμ. Heat treatment bleaches the 785 mμ band and makes the 835 mμ band grow.

VEDENEEVA and RUDNITSKAYA[38] have made absorption spectra measurements, using polarized light, on smoky quartz in its natural state, after bleaching at 251°C for 4 hr, and after complete decoloration at 400°C and X-irradiation. A maximum at 440 mμ was found for both the extraordinary and ordinary rays in all three cases. The absorption is most intense for the extraordinary ray and thus the center is anisotropic. These authors believe the color centers to be due to electrons trapped at deformed regions caused by the presence of foreign cations in the channels lying parallel to the *c*-axis. The non-spherical nature of these regions would give rise to the observed anisotropy.

Neutron irradiated soda-lime glass and fused quartz was studied by MAYER and GUÉRON[39]. They found absorption maxima at 340 mμ, 420 mμ and 650 mμ in the glass, while fused quartz gave bands at 550 mμ, 300 mμ and 220 mμ.

YOKOTA[40] studied both fused and crystalline quartz. For X-irradiated crystalline quartz he found a broad

Table 1

Material and treatment	Wavelengths of absorption maxima (mμ)	Investigator
Smoky quartz	~600	Holden[6]
Smoky quartz	425	Mohler[27]
Smoky, bleached γ-irradiated smoky, γ-irradiated clear quartz	606 500 400 303 263 488 294	Mohler[28,29]
X-irradiated natural quartz	(smoky) 390 (violet) 302	Futagami[30]
Smoky quartz (dichroic maxima)	400 300	Vedeneeva[33]
X-irradiated natural quartz	435–400 < 205	Smakula[34]
X-irradiated natural quartz Neutron irradiated natural quartz (Both cut perpendicular to *c*-axis)	~400 470–450	Forman[35]
Smoky, partially bleached smoky, bleached X-irradiated smoky, X-irradiated clear quartz (All cut perpendicular to *c*-axis)	440 (*e*- and *o*-rays)	Vedeneeva and Rudnitskaya[38]
X-irradiated natural quartz (Cut parallel to *c*-axis)	462	Yokota[40]
Smoky, bleached X-irradiated smoky (Both cut perpendicular to *c*-axis) γ-irradiated natural quartz (Cut parallel to *c*-axis)	625 425	Cohen[43]
X-irradiated synthetic quartz (Cut perpendicular to *c*-axis)	470	Arnold[44]
Neutron irradiated natural quartz X-irradiated natural quartz Neutron irradiated synthetic quartz X-irradiated synthetic quartz Smoky quartz Neutron irradiated bleached smoky quartz (All cut parallel to *c*-axis)	618 468–459 212–208 618 459 214 477 212 477 214 618 468 (structure at 230) 207 618 451 217	Mitchell and Paige[48]
Smoky quartz, X-irradiated natural quartz	455 430 (cut parallel to *c*-axis) 470 455 (cut perpendicular to *c*-axis)	Cohen[53]

VOL. 13 1960

Table 1. (Contd.)

Material and treatment	Wavelengths of absorption maxima (mμ)	Investigator
X-irradiated Ge and B doped synthetic quartz. (Cut parallel to *AT* seed)	625 470 285–280 210	Arnold[63]
Neutron irradiated natural and synthetic quartz	163	Mitchell and Paige[5]
X-irradiated natural and synthetic quartz	173	
Smoky quartz	260 240	Arnold (this paper)
Amethyst quartz	540–530	Holden[6]
Amethyst quartz	546 262	Mohler[27]
Amethyst quartz (Deccan)	920 835 785 550	Bappu[36,37]
Amethyst quartz (Brazil, Colorado, Michigan, N. Carolina)	950 540 340 266 225	Cohen[66]
X- and γ-irradiated fused quartz	300	Choong[31]
X-irradiated fused quartz	540	Vedeneeva[32]
X-irradiated fused quartz	540 305	Smakula[34]
X-irradiated fused quartz	539 302 221	Yokota[39]
Neutron irradiated fused quartz	539 305 238 218	Mitchell and Paige[43]
X-irradiated fused quartz	539 305 238 211	
X-irradiated fused quartz	545 300 220	Dainton and Rowbottom[49]
X-irradiated Vitreosil (High Boron)	539 305 223	Levy and Varley[55]
Neutron irradiated Corning fused silica	242 218	Levy[59]
X-irradiated fused quartz (Amersil) after reducing treatment	540 300 242 214	Cohen[60]
X-irradiated Corning fused silica	214	
Neutron and X-irradiated fused quartz	163	Mitchell and Paige[5]
X-irradiated fused quartz	173	

absorption maximum at 462 mμ. For fused quartz prepared under reducing conditions, he found that X-irradiation gave a purple color and absorption maxima at 539 mμ, 302 mμ and 221 mμ. When the fused quartz was prepared under oxidizing conditions and X-irradiated, no discoloration was produced and the peak at 539 mμ had disappeared. In a second paper, YOKOTA[41] dealt with fused quartz only. He found that fused quartz prepared under oxidizing conditions and then heated at high temperatures in silicon vapor, yielded the 539 mμ peak on irradiation. Fused material which failed to yield the 539 mμ band when prepared under oxidizing conditions could be made to display the peak by re-processing under reducing conditions. YOKOTA concluded that the peak at 539 mμ was caused by the trapping of electrons at oxygen vacancies. NIIRA[42] has attempted to explain the difference in the peak position of the visible absorption maximum in crystalline and fused quartz which YOKOTA believes to be due to the same defect. The argument is based on the difference in energy between the bound states of an electron on a silicon atom adjacent to an oxygen vacancy. He finds agreement with YOKOTA's data if a Si–Si distance of 3·22 Å is used for fused quartz and 3·08 Å for crystalline quartz.

COHEN[43] found natural and bleached X-irradiated quartz to yield absorption maxima at 425 mμ and 625 mμ for orientation either parallel or perpendicular to the optic axis. Natural colorless quartz (not parallel to *c*-axis) also gave the same peaks after γ-irradiation. He believed the similarity between the position of the visible absorption band in amethyst and fused quartz to be evidence that the band in fused quartz is, like that in amethyst, an *F*-band, and that YOKOTA and NIIRA are incorrect in their interpretation of the visible bands in fused and crystalline quartz.

The present author[44] studied synthetic quartz (Clevite Corp. and Bell Laboratories, Inc.) and found an X-ray induced maximum at 470 mμ for sections perpendicular to the *c*-axis. An argument was advanced in support of YOKOTA's interpretation of electron trapping at oxygen vacancies, based on the reduced absorption at 2·92 μ for synthetic quartz as opposed to natural quartz. It was thought that the 2·92 μ band might be a third harmonic of the valence bond oscillation of the oxygen atoms believed by SAKSENA[45] to give rise to the characteristic infra-red absorption at 8·7 μ in quartz. (See, however, SIMON and MCMAHON[46]; MITCHELL and RIGDEN[47].)

MITCHELL and PAIGE[48] measured the absorption of fused, natural and synthetic quartz for both X- and neutron irradiation. Absorption maxima were noted at 618 mμ, 459–468 mμ and 208–214 mμ for natural crystalline quartz samples. In the synthetic samples the maxima were observed at 477 mμ and 212–214 mμ. Smoky quartz gave maxima at 618 mμ, 468 mμ and 207 mμ with some structure noted at 230 mμ. Smoky quartz which was bleached thermally and neutron irradiated gave maxima at 618 mμ, 451 mμ and 218 mμ. Measurements on fused quartz were in agreement with measurements made by YOKOTA.[40,41] These authors also noted the radiation bleaching of the visible bands with neutron doses in excess of $10^{17}/cm^2$, with fused and crystalline quartz.

DAINTON and ROWBOTTOM[49], in a study of the kinetics of the coloration and luminescence of vitreous silica, found absorption maxima at 545 mμ, 300 mμ and 220 mμ. The irradiation was with X- and γ-rays. The centers at 220 mμ and 300 mμ could be interconverted by light of the appropriate wavelength. The 545 mμ band could not be bleached by light in the band but was bleached by light in the 220 mμ band. These authors noted a radiation bleaching effect for the 300 mμ band when irradiation was carried out beyond saturation.

GRIFFITHS *et al.*[50] concluded from paramagnetic resonance measurements of neutron irradiated natural and synthetic quartz and smoky quartz that an impurity with nuclear spin 5/2 is responsible for the smoky color. Al^{27}, Mn^{55} and I^{127} were suggested, with aluminum being the most likely. This paper presents the first real evidence for an impurity as being responsible for the coloration in crystalline quartz. Further experimental work by these authors[51] and the interpretation of their data by O'BRIEN and PRYCE[52] has established aluminum as the impurity responsible for the visible color in quartz. The anisotropic absorption of this center in crystalline smoky and X-irradiated quartz, noted previously by SMAKULA[34] and VEDENEEVA and RUDNITSKAYA[38], was confirmed by COHEN[53]. He, however, found the center to be composed of two peaks which overlapped at saturation values of coloration and appeared to be a single peak. He found the two peaks to be located at 430 mμ and 455 mμ, for quartz cut parallel to the *c*-axis, giving a single peak at 465 mμ at X-ray saturation. For sections cut perpendicular to the *c*-axis, the doublet was given by 455 mμ and 470 mμ, with higher X-ray doses giving a maximum centered about 460 mμ.

Rate of growth curves for the various bands in neutron irradiated crystalline and fused quartz were obtained by DITCHBURN *et al.*[54] It was shown that the ultraviolet bands (called *C*-bands) did not saturate with neutron dose. These bands did saturate with X-ray irradiation. Therefore, the defect responsible for the absorption in irradiated crystalline and fused quartz is present in the unirradiated material and is a product of displacing radiation. Analysis of five crystalline quartz specimens for aluminum content showed the absorption in the *C*- and *A*-(visible) bands to be proportional to the aluminum content but the correlation was poor. The suggestion was made that some aluminum was present interstially.

LEVY and VARLEY[55] studied irradiated fused quartz (Vitreosil) and found the usual bands, i.e. 223 mμ, 305 mμ and 539 mμ. Their spectrographic analysis showed 0·3 per cent by weight of boron. Irradiated glasses of sodium tetraborate composition give a maximum at 539 mμ whose shape was similar to that in the fused quartz. The strong absorption at 243 mμ in the unirradiated glass and a small absorption in the unirradiated fused quartz at the same wavelength was considered further evidence that sodium tetraborate was responsible for the color. (See, however, GARINO-CANINA[56] and COHEN[57] in which evidence is given for germanium as the impurity responsible for the 243 mμ band in unirradiated fused quartz).

VOL. 13 1960

The high resistance to irradiation coloration of Corning fused silica[58] was verified by P. W. LEVY[59] and COHEN[60]. LEVY found that 10^6 r of Co^{60} γ-irradiation produced virtually no absorption in the Corning material from 200 mμ to 1000 mμ, a dose which colors most crystalline and fused quartz. Neutron irradiation developed bands at 218 mμ and 242 mμ but no visible absorption. COHEN worked with fused quartz and Corning fused silica samples which did not show a 242 mμ impurity band before irradiation. This band could be induced in the fused quartz by heat treatment in a reducing environment but not in the Corning material. It was, therefore, concluded that the band was due to a reduced impurity present in the fused quartz but not in the Corning sample. Since the 540 mμ and 300 mμ bands could not be developed in the Corning fused silica it was believed that these bands also were due to impurities.

MITCHELL and PAIGE[61] and COHEN[62] independently studied the anisotropy of the visible color centers in crystalline quartz. MITCHELL and PAIGE found that Al^{3+} substituting for Si^{4+}, with ionization providing a trapped hole which spends most of its time on an adjacent oxygen as proposed by O'BRIEN and PRYCE[52], is not inconsistent with the anisotropy observed. COHEN showed that aluminum was the only impurity in his specimens which could be correlated with the measured absorption coefficients.

The present author made a study[63] of synthetic quartz in which impurities had been deliberately introduced. It was found (for crystals grown on *AT* seeds) that the optical density of the visible and ultraviolet bands was not proportional to the aluminum concentrations present in the lattice. It was suggested that aluminum was present both interstially and substitutionally. Optical bleaching and re-irradiation (X-rays) resulted in faster growth of these bands, but the same saturation values were obtained. This is similar to results obtained by FORMAN[35] with thermal bleaching. A new band near 280 mμ was observed in synthetic quartz containing germanium and boron as an impurity. An argument based on short-range disorder of the region surrounding the impurity was advanced as a possible explanation of the absorption rather than the specific impurity. Growth curves for the bands at 210 mμ and 280 mμ in the germanium doped crystals reached maximum values fairly rapidly and then dropped slowly with continued irradiation. Light in the 280 mμ band caused fading of all the color centers.

STEVELS and KATS[64] have proposed a nomenclature for the description of defects in silicates and have used their system[65] in an extensive investigation of irradiated glasses, fused quartz and natural and synthetic quartz. The general features of previous absorption studies were confirmed with the addition of resolved fine structure bands in irradiated crystalline quartz. (These latter are too numerous for inclusion in Table 1 and the referenced paper should be consulted.)

Amethyst color center studies were extended by COHEN[66] who found bands which had the attributes of color centers at 225 mμ, 266 mμ, 340 mμ, 540 mμ and 950 mμ. The 340 mμ band was found to be anisotropic. The similarity of the 540 mμ band in amethyst and fused quartz and complex glasses is remarked and it was suggested that the same impurity is responsible for the coloration of all.

A model was proposed by MITCHELL and PAIGE[5] which accounts for the well documented bands in irradiated crystalline α-quartz and for the far ultraviolet bands found by these authors and described in their paper. These new bands are at 163 mμ (called the *E*-band) in neutron irradiated α-quartz and at 173 mμ (called the *D*-band) in X-rayed material. The presence of the *D*-band in neutron irradiated specimens is uncertain because of the overlap of the *C*- (218 mμ) and *D*-bands. These bands have counterparts in fused quartz where they are found at approximately the same wavelengths. The *E*-band in the fused material, however, can also be developed by X-rays. The model proposed is that of O'BRIEN and PRYCE[52] for the visible absorption maxima (A_1- and A_2-bands), i.e. a hole, left by an electron ionized off a substituted Al^{3+} ion, which is shared with an oxygen neighbor. An oxygen vacancy with a trapped electron is suggested for the *C*-band with the *E*-band being a nearby interstitial oxygen with a trapped hole. Neutron induced defects can be annealed at 950°C but such treatment results in increased absorption in the *A*-bands on re-irradiation. It was suggested that aluminum present interstitially was made to diffuse into vacant silicon sites to account for this enhanced absorption.

The work of GARINO-CANINA[56] and of COHEN[57] has established the 242 mμ band found in some unirradiated fused quartz and silica as being due to germanium in a reduced state. COHEN[57] has also shown that a color center is developed at this same wavelength by neutron irradiation of fused silica.

(b) *Thermoluminescence*

The concomitant luminescence with heating of smoky quartz has been known for hundreds of years. Early studies[14,16–19] established that smoky quartz begins to bleach (and luminesce) in the neighborhood of 200°C. Luminescence usually ceased after 300°C. Amethyst requires higher temperatures ($\sim$ 400°C) for de-coloration. The apparent similarity between natural colors and the color induced by γ-irradiation in de-colorized smoky quartz and amethyst quartz was noted.[13,14] These artificial colors could also be removed by heating. Other references to work before about 1924 can be found in HOLDEN[6].

FRONDEL[67], in experiments with X-irradiated oscillator plates, noted that the thermoluminescence of irradiated quartz begins below 100°C while in natural smoky quartz the threshold for luminescence was near 180°C. The observation that smoky quartz is a feeble emitter when heated and that γ-irradiated bleached smoky quartz was intensely thermoluminescent was made by DANIELS and SAUNDERS[68]. These workers found that γ-irradiated material luminesced most strongly at 180°C with a subsidiary maxima at 290°C.

CHENTSOVA and VEDENEEVA[69] studied the rate of decolorization of smoky quartz with temperature and found

the rate to become discontinuous at 251°C. They also note a faster decoloration rate in artificially colored quartz and suggest two different kinds of centers for the visible absorption in quartz with the concentration and stability of each being different in naturally colored material.

The thermoluminescence of quartz and fused quartz colored to saturation values (X-rays) was measured by YOKOTA[70]. He prepared fused quartz in reducing and oxidizing conditions and measured the glow curves. As noted above, YOKOTA[40,41] has shown that the visible absorption band at about 550 mμ is absent in fused quartz prepared under oxidizing conditions. The main emission maximum is at about 320°C for quartz fused in a reducing atmosphere and at about 120°C in the oxidizing one. Subsidiary maxima are present in both cases. Short periods of decay at temperatures near the maximum emission temperature shifts the emission to higher temperatures, thus showing that the fused quartz maxima are not caused by transitions within discrete levels but are composed of groups of discrete levels.

LAUTOUT[71] has studied the thermoluminescence of fused quartz. LAUTOUT[72] has also shown the inhomogeneity of fused quartz with respect to the defects responsible for emission after irradiation by photographing the surface of the fused quartz in its own emitted light. She, however, did not find[73] that crystalline quartz fluoresced, phosphoresced or changed in color with 5×10^5 r X-irradiation. This is in conflict with the experience of most investigators.

References to the thermoluminescence of glasses and fused quartz can be found in the review articles by SUN and KREIDL[74] and by OTLEY[75].

EXPERIMENTAL

Measurements of the thermoluminescence of natural and synthetic quartz, fused silica and a soda-lime glass have been made from room temperature to about 380°C. The apparatus used was designed by Dr. P. PIZZOLATO* and modified by R. J. GINTHER of this laboratory and will not be described in detail here. The experimental procedure consists of the measurement of the emission from solid samples by means of the amplified signal from a photomultiplier tube while the temperature of the sample is varied at a uniform rate. Temperature and light intensity are recorded simultaneously on an x–y recorder. The light striking the photomultiplier tube was first passed through a suitable filter to eliminate wavelengths above about 700 mμ. The rate of heating was about 17°C/min.

The specimens used in this investigation were usually 2–4 mm in thickness, highly polished on both surfaces and about 1/4 in. × 1/4 in. in cross-section. Irradiation was carried out variously with X-rays (45 kVp at 10 mA), γ-rays (Co^{60} source) and ultraviolet light (usually from a low pressure mercury arc which is rich in 2537 Å radiation). The sensitivity of thermoluminescence measurements will be noted from the fact that very short exposures to 2537 Å light are efficient in producing thermoluminescence in crystalline quartz, whereas changes in absorption are not noted even after several days of irradiation.

* Now at the Argonne National Laboratory, Lemont, Ill.

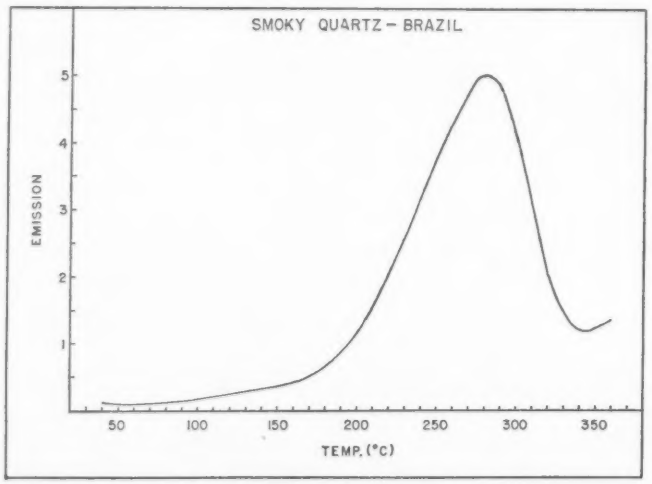


FIG. 1. Thermoluminescence (arbitrary units) vs. temperature for Brazilian smoky quartz.

Fig. 1. is a plot of emitted light intensity vs. temperature for a specimen of Brazilian smoky quartz. The emission is feeble in comparison with artificially colored quartz, a fact which has been previously noted.[68] The peak emission for this specimen is at about 245°C while for another sample it was at about 280°C.

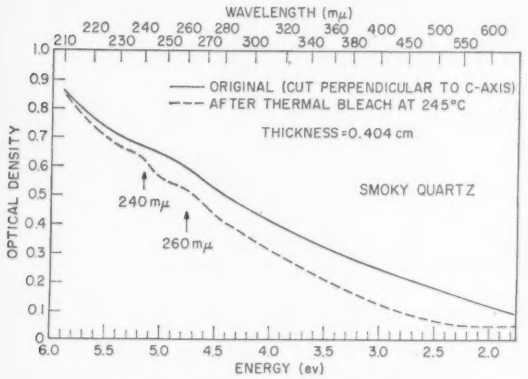


FIG. 2. Optical density vs. absorbed energy (eV) for Brazilian smoky quartz.
——— before thermal bleach
– – – – – after thermal bleach at 245°C

It was noted that the smoky quartz specimen still retained some color after the thermoluminescence run was completed. The absorption of another sample (from the same quartz which gave maximum emission at 245°C) was measured before and after heat treatment at 245°C for 24 hr. The results are shown in Fig. 2. The difference curve shows preferential bleaching in the vicinity of 420 mμ, with subsidiary maxima at about 235 mμ and 260 mμ. The 235 mμ maximum can also be observed as an inflection in the absorption curves both before and after thermal bleaching.

A different smoky quartz sample which did not display a maximum at 235 mμ at room temperature was measured at 77°K. Fig. 3 shows the absorption for both room temperature and 77°K. There is a slight maximum at about 240 mμ at the lower temperature.

Another specimen from Graham Co., Arizona, exhibited the thermoluminescence spectra shown in Fig. 4. This specimen would be described as "clear" quartz on the basis of visual inspection and spectrophotometric absorption measurements. It is seen, however, to have a peak luminescence at 245°C, as in the smoky quartz samples, with a small maximum in the vicinity of 330°C.

With artificially colored material the thermoluminescence emission is more intense and the maximum emission occurs at higher temperatures than in naturally colored quartz. It is possible to obtain well defined thermoluminescence peaks with X-irradiation times of the order of seconds as can

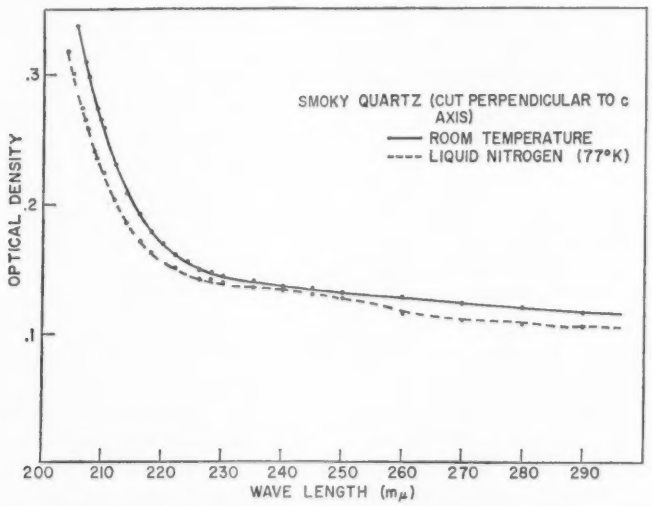


FIG. 3. Optical density vs. wavelength (mμ) for Brazilian smoky quartz at room temperature and at 77°K.

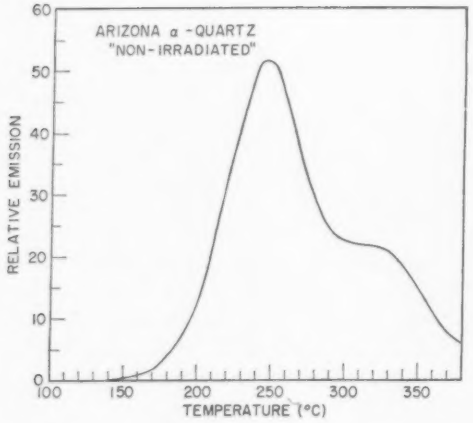


FIG. 4. Thermoluminescence (arbitrary units) vs. temperature for Arizona (Graham Co.) quartz.

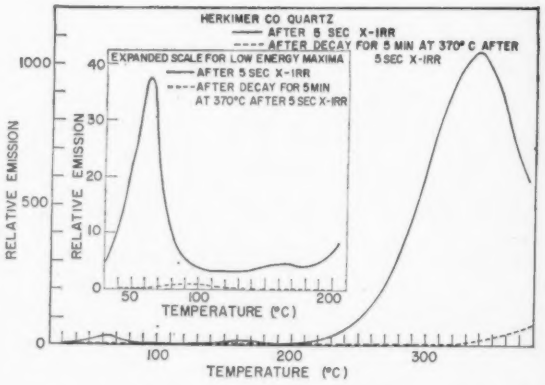

HERKIMER CO QUARTZ
AFTER 5 SEC X-IRR
AFTER DECAY FOR 5 MIN AT 370° C AFTER 5 SEC X-IRR
EXPANDED SCALE FOR LOW ENERGY MAXIMA
AFTER 5 SEC X-IRR
AFTER DECAY FOR 5 MIN AT 370°C AFTER 5 SEC X-IRR
RELATIVE EMISSION
TEMPERATURE (°C)

FIG. 5. Thermoluminescence (arbitrary un ts) vs. temperature for New York (Herkimer Co.) quartz.
———— after 5 sec X-irradiation
– – – – – after decay for 5 min at 370°C (after 5 sec X-irradiation).

be seen in Fig. 5, for a sample of quartz from Herkimer Co., New York. This crystal was X-irradiated for 5 sec and the thermoluminescence measurement was made 5 min after the irradiation. The measurement shows three maxima, two small ones at 63°C and 164°C and a large one at 338°C. The low temperature maximum is unstable, falling rapidly at first and then more slowly but is still measurable several hours after irradiation. A crystal which has been measured over the range from room temperature to 380°C and which shows no thermoluminescence when run repeatedly over this range and no absorption maxima, can be made to yield an intense thermoluminescence peak at 340°C by a short exposure (~ 1 sec) to X-irradiation or ultraviolet light (2537 Å). This can be repeated time after time and can be eliminated by a heat treatment of several hours at temperatures in the neighborhood of 550°C. This bears a resemblance to the "re-activation" found by LAUTOUT[72] for fused quartz which had been X-irradiated and then made to fluoresce after phosphorescence decay, by application of ultraviolet light. Apparently, in the present case, the original thermal bleach at temperatures near 400°C is inadequate to return all electrons to their ground state and a burst of energy removes them from metastable levels.

As a first attempt to investigate the nature of the center responsible for the emission at 338°C, the crystal was again X-irradiated for 5 sec (after a long bleach at 500°C) and then allowed to decay in an oven set for 370°C for five minutes. It was then brought quickly to room temperature and the thermoluminescence was measured after the same total time lapse as in the first measurement. The result is shown by the broken curve in Fig. 5. The high temperature maximum appears to have shifted to a higher temperature beyond the range of measurement. Similarly the low temperature emission, as seen in the inset, has shifted to a new peak position at about 95°C. The other low temperature peak is not measurable after the treatment.

In Fig. 6 are shown the results for a sample from the same quartz in which the X-irradiation time was for one minute and the decay for 2 min at 330°C, after high temperature bleach and re-irradiation for 1 min. Here the shift of the emission remains within the measurement range, the high temperature maximum shifting from 338°C to 346°C and 60°C to 84°C for the low temperature maximum.

The shift in peak position with decay at temperatures near the maximum emission temperature indicates that the emission at 338°C is made up of the emission from many discrete trapping levels rather than from one level. If the latter had been the case there would have been no shift in peak position but only a diminution of intensity. This also explains why it is necessary to go to much higher temperatures in order to empty all traps after apparently exhausting the luminescence at

VOL. 13 1960

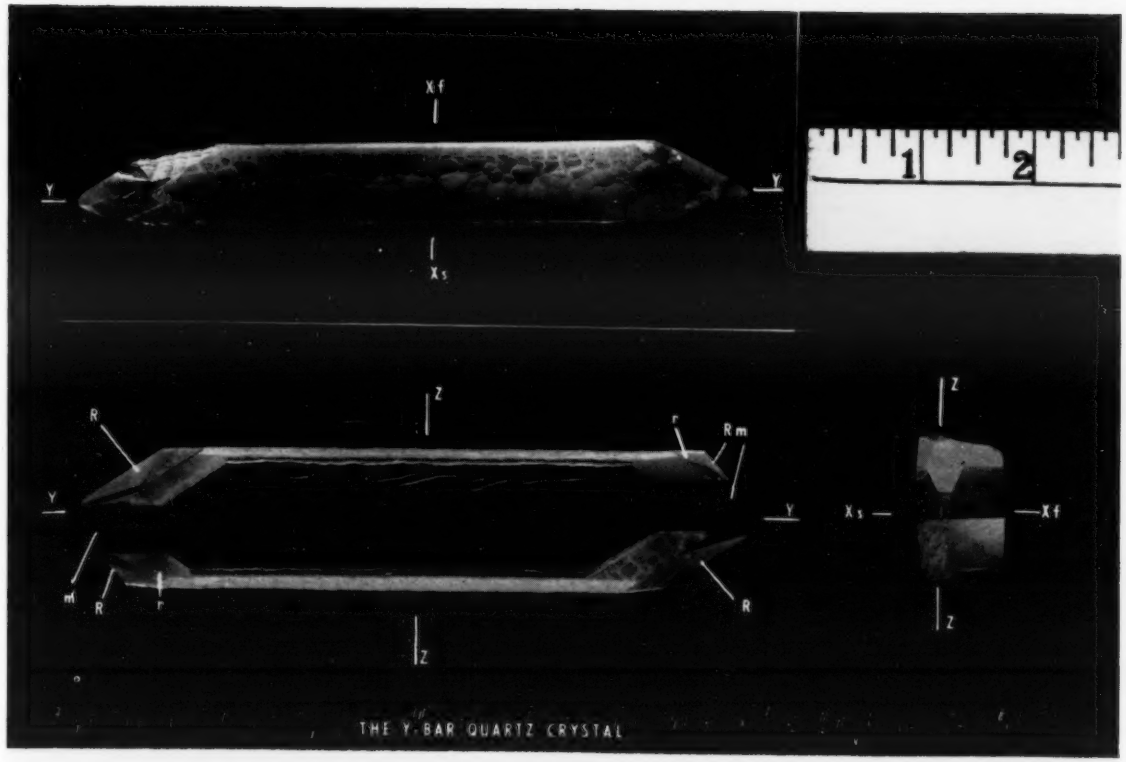


FIG. 8. *Y*-bar synthetic quartz (Courtesy of Dr. DANFORTH R. HALE, Clevite Research Center, Cleveland, Ohio).

[*facing p.* 314

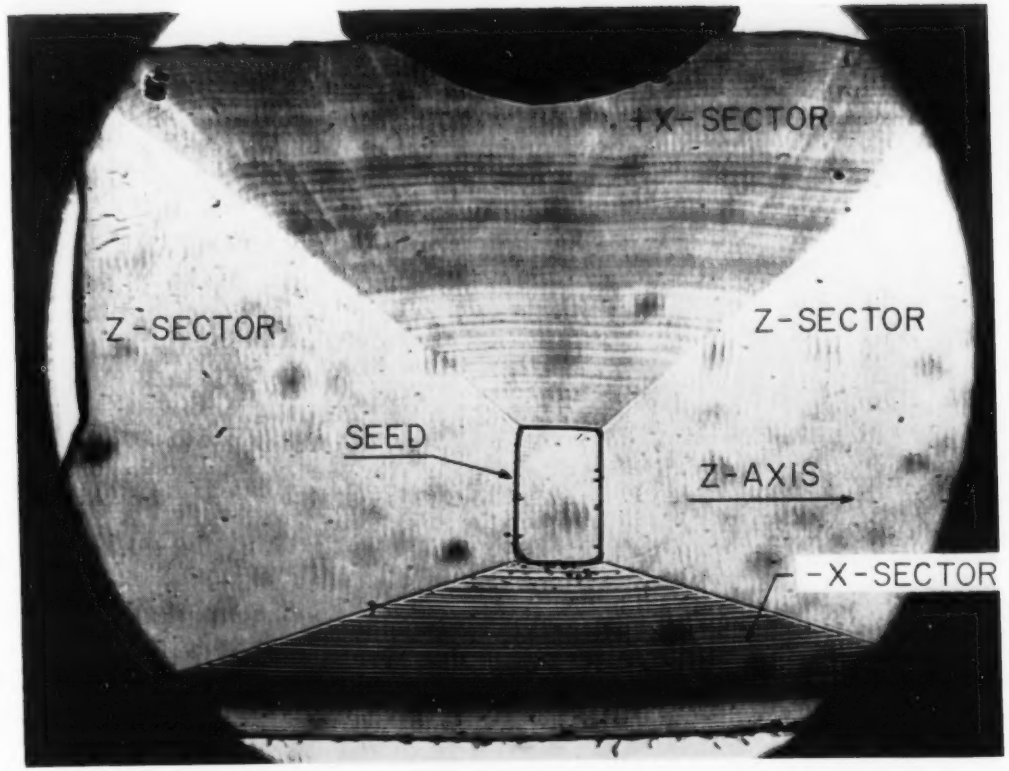


FIG. 9. γ-irradiated sections of Y-bar quartz (cut perpendicular to Y-axis).

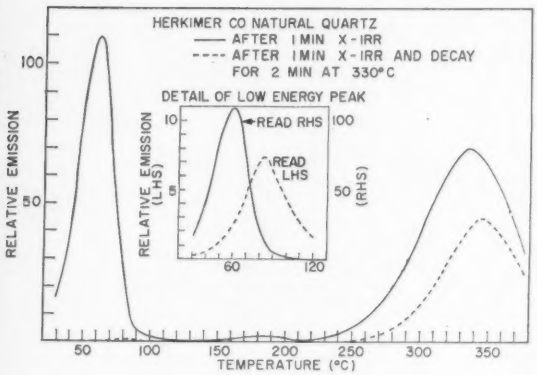


FIG. 6. Thermoluminescence (arbitrary units) vs. temperature for New York (Herkimer Co.) quartz.
——— after 1 min X-irradiation
- - - - - after 1 min X-irradiation and decay for 2 min at 330°C.

380°C. The electrons in these higher levels are available to be kicked into intermediate trapping centers by short bursts of X-rays or of 2537 Å light.

The effect of annealing at high temperatures prior to irradiation is shown in Fig. 7 for a specimen of α-quartz from Crystal Springs, Arkansas. The solid curve is for the original sample after γ-irradiation. The γ-dose was of the order of 10^7 r and irradiation was carried out some weeks before

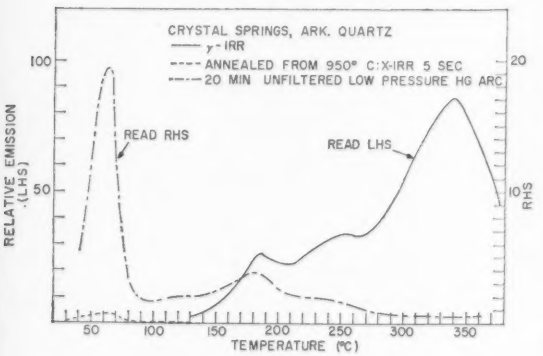


FIG. 7. Thermoluminescence (arbitrary units) vs. temperature for Arkansas (Crystal Springs) quartz.
——— after 10^7 r γ-irradiation
- - - - - after anneal from 950°C and five sec γ-irradiation
— - — - after 20 min irradiation with unfiltered light from low pressure Hg arc.

the thermoluminescence measurements were made. Peak temperature emissions are at 185°C, 255°C and 340°C. The crystal was then held at 950°C overnight and brought to room temperature over a 12-hr period. It was then X-irradiated for 5 sec and the thermoluminescence measurements made. The result is very feeble emission centered about 62°C and no emission at higher temperatures. A 20-min irradiation with an unfiltered low pressure mercury arc results in more intense emission at 62°C and a development of the peaks at 185°C and 255°C, but with negligible emission at 340°C. This crystal did develop luminescence of low intensity at 340°C after about 10^6 r of γ-irradiation.

It can hardly be doubted that the emission centered about 340°C is associated with the optical absorption maximum at about 210 mμ in irradiated crystalline quartz. The visible and ultraviolet absorption maxima diminish rapidly at this temperature and the onset of bleaching coincides with the observed rise in thermoluminescence intensity. MITCHELL and PAIGE have shown[5] that treatment of neutron irradiated crystalline quartz by heating at 950°C and slowly cooling anneals out the defects responsible for the absorption at 210 mμ. In the present case, similar treatment is effective in reducing the thermoluminescence intensity at 340°C. Some crystals which were particularly resistant to coloration also failed to show high temperature emission.

Synthetic quartz has also been examined in this investigation. The α-quartz grown by the Clevite Corp., and known as Y-bar quartz,* is shown in Fig. 8. A γ-irradiated section of such a crystal, cut perpendicular to the Y-axis, is shown in Fig. 9. The distinctive features of this quartz have been discussed by the author[63,76] and COHEN (this symposium). This slice of quartz was cut into four sections, excluding the seed, and thermoluminescence measurements were made. The γ-irradiation dose was of the order of 10^7 r. The results for the Z-sectors are shown in Fig. 10 and for the X-sectors in Fig. 11. The emission for the Z-sectors peaks at 318°C and 330°C and in the neighborhood of 200°C. The X-sectors yield emission that peaks at higher temperatures, i.e. at 360°C for the $-X$-sector and 380°C for the $+X$-sector. The lower

* Patent application U.S. Serial No. 459,052 of 9/29/54.

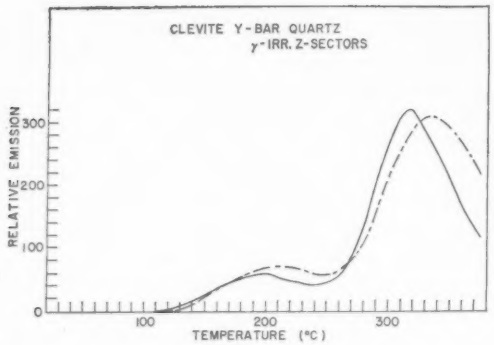


FIG. 10. Thermoluminescence (arbitrary units) vs. temperature for γ-irradiated Z-sectors of Y-bar quartz section.

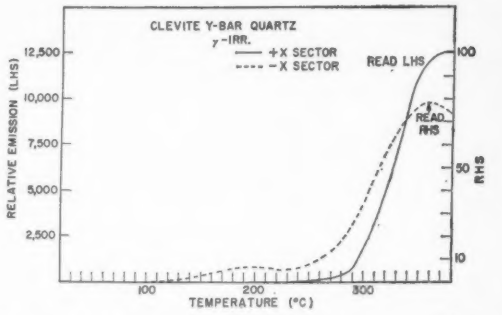


FIG. 11. Thermoluminescence (arbitrary units) vs. temperature for γ-irradiated X-sectors of Y-bar quartz section.

———— $+X$-sector
– – – – – $-X$-sector.

temperature maxima are at about 200°C. It will be noted that the luminescent intensity is much less for the $-X$-sector than for the $+X$-sector. COHEN (this symposium) has given evidence that aluminum is present interstially in the Z-sectors and both interstitially and substitutionally in the X-sectors.

Synthetic quartz grown on an AT-cut seed by the Signal Corps Engineering Laboratories and intentionally doped with aluminum was irradiated for 1 hr with light predominantly of 2537 Å wavelength. The result is shown in Fig. 12. The low temperature maxima are at about 62°C and 110°C, while the high temperature emission is centered about 230°C. This crystal had previously been found not to give as much coloration after irradiation as one would expect and the suggestion has been made that the aluminum was present both interstitially and substitutionally.[59] The ratio of the intensity of the high temperature emission to

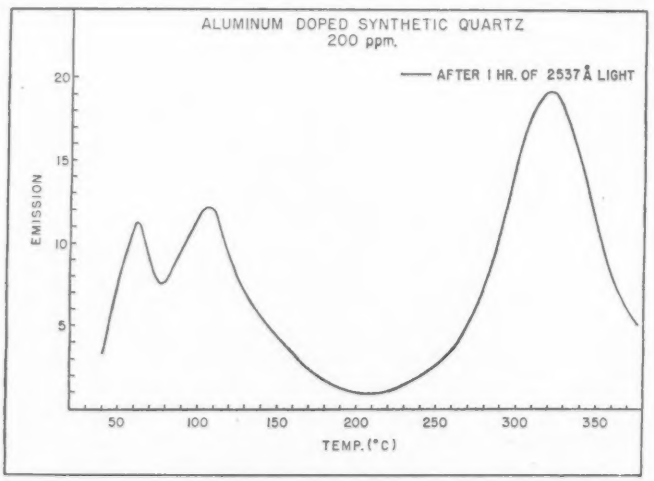


FIG. 12. Thermoluminescence (arbitrary units) vs. temperature of aluminum (200 p.p.m.) doped synthetic quartz after one hour irradiation with predominantly 2537 Å light.

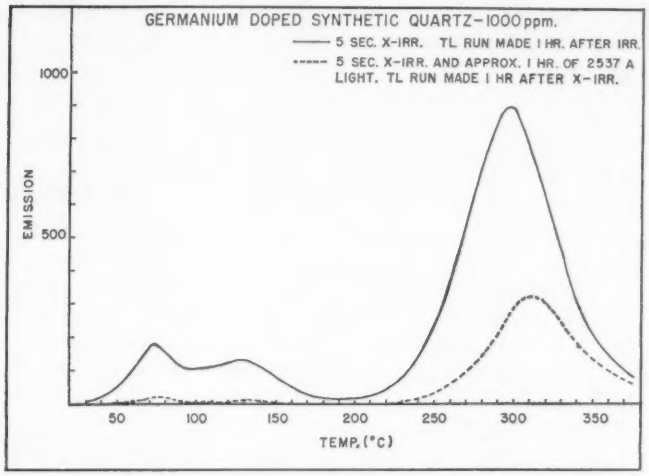


FIG. 13. Thermoluminescence (arbitrary units) vs. temperature for germanium (1000 ppm) doped synthetic quartz.
———— 1 hr after 5 sec X-irradiation
- - - - - 1 hr after 5 sec X-irradiation and approximately 1 hr bleach with 2537 Å light.

that at the lower temperatures is reduced when irradiating with 2537 Å light because of the simultaneous optical bleaching that occurs.

Germanium doped synthetic quartz is distinguished from aluminum doped quartz by differences in the positions of the emission maxima. This is shown in Fig. 13. The peaks appear at 73°C, 128°C and about 290°C. In view of the fact that

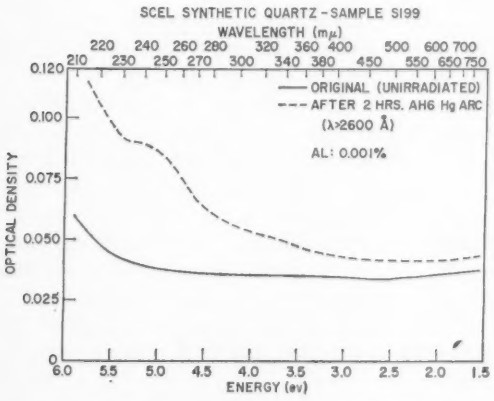


FIG. 14. Optical density vs. absorbed energy (eV) for germanium (1000 p.p.m.) doped synthetic quartz.
———— after 1 hr X-irradiation
- - - - - after thermal bleach at 125°C for 10 min.

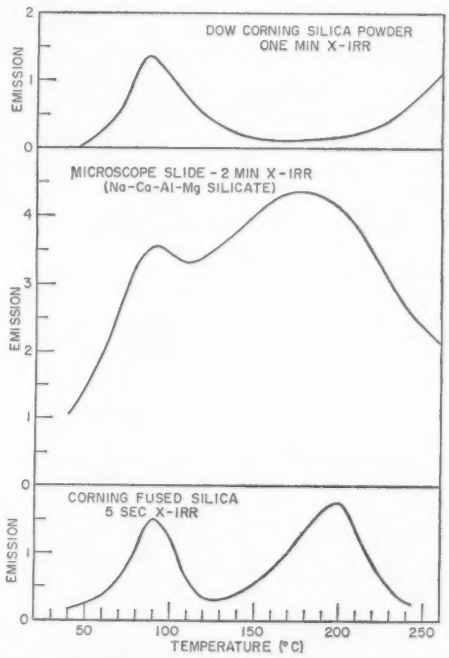

DOW CORNING SILICA POWDER ONE MIN X-IRR
MICROSCOPE SLIDE - 2 MIN X-IRR (Na-Ca-Al-Mg SILICATE)
CORNING FUSED SILICA 5 SEC X-IRR
EMISSION
50 100 150 200 250
TEMPERATURE (°C)

FIG. 15. Thermoluminescence (arbitrary units) vs. temperature for X-irradiated: (a) Dow Corning silica powder; (b) glass microscope slide; (c) Corning fused silica.

germanium doped quartz contains an additional absorption maximum at about 285 mμ,[63] which can be bleached by light of this wavelength, the crystal was re-irradiated for 5 sec and was subjected to the light of a low pressure mercury arc passed by a Corning 9863 filter. This passes the 2537 Å line but cuts off light of wavelengths below 2300 Å. The result is as indicated by the broken curve in Fig. 11. The emission peaks at the lower temperatures are reduced to about 10 per cent of their value on the previous irradiation. The high temperature peak is also reduced, because irradiation in the 285 mμ band in germanium doped quartz bleaches the visible as well as the ultraviolet bands. Bleaching thermally at 125°C to remove the lowest two emission peaks after 1 hr X-irradiation gives the results shown in Fig. 14. The 280 mμ band has decreased as well as the visible absorption bands. The 220 mμ band, however, has increased.

The results (Fig. 15) for Corning fused silica, Dow Corning silica powder and glass are rudimentary but indicate that the emission at about 90°C is caused by a defect common to all three. No high temperature emission has been observed for the moderate irradiation doses given to Corning fused silica.

The effectiveness of ultraviolet irradiation in producing thermoluminescence has led to attempts to bring about color center absorption by this means. The result of irradiation with a General Electric AH6 mercury arc on a synthetic quartz sample is shown in Fig. 16. A glass filter was used to cut off light below 260 mμ. A small absorption peak at about 240 mμ is produced.

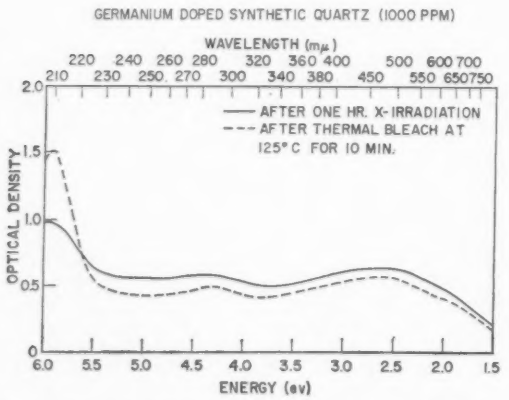

GERMANIUM DOPED SYNTHETIC QUARTZ (1000 PPM)
WAVELENGTH (mμ)
220 240 260 280 320 360 400 500 600 700
210 230 250 270 300 340 380 450 550 650 750
AFTER ONE HR. X-IRRADIATION
AFTER THERMAL BLEACH AT 125°C FOR 10 MIN.
OPTICAL DENSITY
2.0
1.5
1.0
0.5
0
6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5
ENERGY (ev)

FIG. 16. Optical density vs. absorbed energy (eV) for Signal Corps. Engineering Laboratory synthetic quartz (S199).

———— before irradiation

– – – – – after 2 hr irradiation with light from AH6 lamp and glass filter.

DISCUSSION

The color center absorption maximum centered about 210 mμ in irradiated crystalline quartz is clearly not a single center, but is composed of a group of discrete levels. MITCHELL and PAIGE[5] observed a band at about 230 mμ by preferential bleaching of neutron irradiated crystalline quartz. They suggested that, for their model, this absorption arises from an oxygen vacancy–oxygen interstitial pair formed by collisions of just enough energy to cause displacement. They believe the separation of the pair is such that the energy level of the *C*-band (210 mμ) is modified to give absorptions at 230 mμ. It is probable, based on the results of the present investigation, that the energy level of an isolated oxygen vacancy and a trapped electron is modified in a continuous manner by a distribution of vacancy–interstitial separations. LEVY (this symposium) has resolved the absorption in the ultraviolet region, for Corning fused silica, into several discrete bands. Also, in the present paper, bands at 240 mμ and 260 mμ have been observed in smoky quartz and at 240 mμ in ultraviolet irradiated synthetic quartz. The fact that natural smoky quartz thermoluminesces at a much lower temperature than does freshly irradiated material, suggests that the centers responsible for the emission in smoky quartz are the most stable configuration of oxygen vacancy–interstitial pairs. The Arizona quartz (Fig. 4) would seem to be an intermediate case where emission of less stable centers at higher temperatures is still apparent. It might, perhaps, be possible to obtain ideas as to the geologic time lapse since irradiation of smoky quartz by examination of its thermoluminescence emission.

The fluorescence and phosphorescence of X-irradiated crystalline quartz is caused by the emission from the low temperature peaks. These also shift with decay at higher temperatures (Figs. 5 and 6) and are probably metastable states of the centers responsible for the high temperature emission.

The emission of freshly irradiated natural and synthetic quartz and aluminum doped synthetic quartz appears to be the same, especially with respect to the positions of the low temperature maxima. The addition of germanium alters the position of all peaks and one can distinguish germanium-bearing quartz by its thermoluminescence. The differences which exist between aluminum-bearing quartz, on the one hand, and germanium-bearing quartz, also containing aluminum, on the other, suggest that aluminum and germanium can replace each other in the lattice and that the emissions observed are for transitions of an electron bound to an aluminum or germanium ion neighboring an oxygen vacancy.

The work on thermoluminescence of *Y*-bar quartz is inconclusive. There is a greater tendency for emission to occur at lower temperatures in *Z*-sectors than in *X*-sectors and this may have some bearing as to the role of aluminum in the lattice, i.e. whether present interstitially or substitutionally.

Further work is contemplated on all phases of the work reported here.

VOL. 13 1960

Acknowledgements—The author wishes to thank Mr. J. M. STANLEY of the Army Signal Engineering Laboratories, Ft. Monmouth, N. J. and Dr. D. R. HALE, Clevite Research Center, Cleveland, Ohio for the specimens of synthetic quartz used in this investigation.

REFERENCES

1. RANDALL J. T. and WILKINS M. H. F., *Proc. Roy. Soc.* A**184**, 366 (1945).
2. RANDALL J. T. and WILKINS M. H. F., *Proc. Roy. Soc.* A**184**, 390 (1945).
3. RANDALL J. T. and WILKINS M. H. F., *Proc. Roy. Soc.* A**184**, 347 (1945).
4. GARLICK G. F. J. and WILKINS M. H. F., *Proc. Roy. Soc.* A**184**, 408 (1945).
5. MITCHELL E. W. J. and PAIGE E. G. S., *Phil. Mag.* **1**, 1085 (1956).
6. HOLDEN E. F., *Amer. Min.* **10**, 203 (1925).
7. WEINSCHENK E., *Z. anorg. Chem.* **12**, 375 (1896).
8. NABL A., *Mh. Chem.* **20**, 272 (1899).
9. NABL A., *Tschedische Mineralogische Mitterlg.*, **19**, 273 (1900).
10. VON KRAATZ-KOSCHLAN K. and WÖHLER L., *Z. Kristallogr.* **33**, 632 (1900); (Abstract).
11. SPEZIA G., *Z. Kristallogr.* **33**, 636 (1900) (Abstract).
12. KOENIGSBERGER J., *Z. Kristallogr.* **33**, 637 (1900) (Abstract).
13. BERTHELOT M., *C.R. Acad. Sci., Paris* **143**, 477 (1906).
14. GOLDSCHMIDT V. M., *Forh. Vidensk.-Selsk. Krist.* No. 5 (1906).
15. PHILLIPS C. E. S., *Nature, Lond.* **77**, 535 (1908).
16. SIMON K., *Neues Jb. Miner.* **26**, 249 (1909).
17. BRAUNS R., *Zbl. Miner. Geol. Pälaont.* Jg. 1909, 721 (1909).
18. WATSON T. L. and BEARD R. E., *Proc. U.S. Nat. Mus.* **53**, 553 (1917).
19. NEWBERRY E. and LIPTON H., *Nature, Lond.* **101**, 198 (1918).
20. WILD G. O. and LIESEGANG R. E., *Zb. Miner. Geol. Pälaont.* Jg. 1922, 481 (1922).
21. HOLDEN E. F., *Amer. Min.* **8**, 117 (1923).
22. WILD G. O. and LIESEGANG R. E., *Zb. Miner. Geol. Pälaont.* Jg. 1923, 385 (1923).
23. HOLDEN E. F., *Amer. Min.* **9**, 75, 101 (1924).
24. VANZETTI B. L., *Gazz. chim. ital.* **54**, 95 (1924).
25. WILD G. O. and LIESEGANG R. E., *Zb. Miner. Geol. Pälaont.* Jg. 1923, 737 (1923).
26. HOLDEN E. F., *Amer. Min.* **10**, 127 (1925).
27. MOHLER N. M., *Amer. Min.* **16**, 300 (1931).
28. MOHLER N. M., *Phys. Rev.* **45**, 743 (1933).
29. MOHLER N. M., *Amer. Min.* **21**, 258 (1936).
30. FUTAGAMI T., *Proc. Phys.-Math. Soc. Jap.* **20**, 458 (1938).
31. CHOONG S. P., *Proc. Phys. Soc. London* **57**, 49 (1945).
32. VEDENEEVA N. E., *Trav. Lab. crist. Acad. Sci. URSS* 1940, No. 2 107 (1940).
33. VEDENEEVA N. E., *Trav. Lab. crist. Acad. Sci. URSS* 1940, No. 2 87 (1940).
34. SMAKULA A., *J. Opt. Soc. Amer.* **40**, 266 (1950).
35. FORMAN G., *J. Opt. Soc. Amer.* **41**, 377 (1951).
36. BAPPU M. K. V., *Indian J. Phys.* **26**, 1 (1952).
37. BAPPU M. K. V., *Indian J. Phys.* **27**, 385 (1953).
38. VEDENEEVA N. E. and RUDNITSKAYA E. S., *C.R. Acad. Sci. URSS* **87**, 361 (1952); NRL Translation No. 549.
39. MAYER G. and GUÉRON J., *J. Chim. Phys.* **49**, 204 (1952).
40. YOKOTA R., *J. Phys. Soc., Jap.* **7**, 222 (1952).
41. YOKOTA R., *J. Phys. Soc. Jap.* **7**, 316 (1952).
42. NIIRA K., *Busseiron Kenkyu* **50**, 17 (1952); *Chem. Abstr.* **46**, 7421c (1952).
43. COHEN A. J., *J. Chem. Phys.* **22**, 570 (1954).
44. ARNOLD G. W. Jr., *J. Chem. Phys.* **22**, 1259 (1954).
45. SAKSENA G., *Proc. Ind. Acad. Sci.* A**12**, 93 (1940).
46. SIMON I. and MCMAHON H. O., *J. Chem. Phys.* **21**, 23 (1953).
47. MITCHELL E. W. J. and RIGDEN J. D., *Phil. Mag.* **2**, 941 (1957).
48. MITCHELL E. W. J. and PAIGE E. G. S., *Proc. Phys. Soc.* **67**, 262 (1954).
49. DAINTON E. S. and ROWBOTTOM J., *Trans. Faraday Soc.* **50**, 480 (1954).
50. GRIFFITHS J. H. E., OWEN J. and WARD I. M., *Nature, Lond.* **173**, 439 (1954).
51. GRIFFITHS J. H. E., OWEN J. and WARD I. M., *Rep. Bristol Conf. on Defects in Crystalline Solids* p. 81. Physical Society, London (1955).
52. O'BRIEN M. C. M. and PRYCE H. H. L., *Rep. Bristol Conf. on Defects in Crystalline Solids* p. 88. Physical Society, London (1955).

53. COHEN A. J., *J. Chem. Phys.* **23**, 589 (1954).
54. DITCHBURN R. W., MITCHELL E. W. J., PAIGE E. G. S., CUSTERS J. F., DYER H. B. and CLARK C. D., *Rep. Bristol Conf. on Defects in Crystalline Solids* p. 92. Physical Society, London (1955).
55. LEVY M. and VARLEY J. H. O., *Proc. Phys. Soc. Lond.* B**58**, 223 (1955).
56. GARINO-CANINA V., *C.R. Acad. Sci., Paris.* **242**, 1982 (1956).
57. COHEN A. J., *Phys. Rev.* **105**, 1151 (1957).
58. GALE A. J. and BICKFORD F. A., *Nucleonics* **11**, 48 (1953).
59. LEVY P. W., *J. Chem. Phys.* **23**, 764 (1955).
60. COHEN A. J., *J. Chem. Phys.* **23**, 765 (1955).
61. MITCHELL E. W. J. and PAIGE E. G. S., *Phil. Mag.* **46**, 1353 (1955).
62. COHEN A. J., *J. Chem. Phys.* **25**, 908 (1956).
63. ARNOLD G. W. Jr., *Proc. 10th Ann. Symp. on Frequency Control, SCEL, Ft. Monmouth, N.J.*, p. 61 (1956).
64. STEVELS J. M. and KATS A., *Philips Res. Rep.* **11** 103 (1956)
65. KATS A. and STEVELS J. M., *Philips Res. Rep.* **11**, 115 (1956).
66. COHEN A. J., *Amer. Min.* **41**, 874 (1956).
67. FRONDEL C., *Amer. Min.* **30**, 432 (1945).
68. DANIELS F. and SAUNDERS G. F., *The Thermoluminescence of Crystals*, AECU-1583.
69. CHENTSOVA L. G. and VEDENEEVA N. E., *Dokl. Akad. Nauk. SSSR* **60**, 649 (1948).
70. YOKOTA R., *Phys. Rev.* **91**, 1013 (1953).
71. LAUTOUT M., *J. Chim. phys.* **52**, 259 (1955).
72. LAUTOUT M., *J. Chim. phys.* **52**, 267 (1955).
73. LAUTOUT M., *C.R. Acad. Sci., Paris* **238**, 2409 (1954).
74. SUN K. and KREIDL N. J., *Glass Ind.* **33**, 511 (1952).
75. OTLEY K. O., *Glass Ind.* **33**, 24 (1952).
76. ARNOLD G. W. Jr., *Proc. 11th Symp. on Frequency Control, SCEL, Ft. Monmouth, N.J.*, p. 112 (1956).

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 13. pp. 321–325. Printed in Great Britain.

SUBSTITUTIONAL AND INTERSTITIAL ALUMINUM IMPURITY IN QUARTZ, STRUCTURE AND COLOR CENTER INTERRELATIONSHIPS*

A. J. COHEN†

Mellon Institute, Pittsburgh 13, Pennsylvania

(*Received* 27 *May* 1959)

VOL. 13 1960

Abstract—That color centers in the 460 and 625 mμ regions of the quartz spectrum are related to substitutional aluminum impurity is now well established. Indirect evidence has indicated that interstitial aluminum may appear in quartz. BUERGER has discussed the stuffing of quartz by β-eucryptite, $LiAlSiO_4$, with cell dimensions a little more than double those of β-quartz. Recently, other workers have found that β-spodumene, $LiAl(SiO_3)_2$, is also isomorphous with β-quartz.

General Electric Co. synthetic quartz grown from a *Z*-seed was found in this laboratory to contain in ppm Si atoms: 290 Al; 260 Li; 8 Fe; and 131 Na. X-irradiation for long periods gave no change in the optical absorption of this quartz down to 1850 Å. It was concluded that the aluminum present was interstitial. An attempt to precipitate the interstitial aluminum and lithium was successful. This separate phase was found to "decorate" the macromosaics and to indicate the manner of growth of synthetic quartz on various seed orientations.

Small regions of rhombohedral growth on General Electric Co. *Z*-seed crystals did color due to the substitutional aluminum content. There was no precipitation of interstitial aluminum in these regions. The increase of the spin–lattice relaxation time of the Si^{29} nucleus after precipitation indicates that the paramagnetic trace impurity in *Z*-seed crystals is also interstitial. Precipitation experiments on *Y*-bar synthetic quartz coupled with X-irradiation indicate that the trace of aluminum present in the *Z*-zones is interstitial and that the major portion of the aluminum in the X_{fast} zone is substitutional, while the aluminum in the X_{slow} zone is both substitutional and interstitial. The color banding in X-rayed *Y*-bar crystals is discussed as well as the dependence of the intensity of the smoky color (due to the aluminum color center system) on the crystallographic orientation of the growing zone and how this can be used to identify the different vicinal faces and their "birth" during growth of the crystal. Mention is made of the effects of germanium (IV) ion on the substitutional aluminum content of the *Z*-zones.

INTRODUCTION

THE nature of the color center system in α-quartz related to substitutional aluminum impurity is well established.[1–8] It consists of two absorption bands at 460 and 625 mμ. In naturally colored smoky quartz the more intense band in the 460 mμ region is apparently shifted to a shorter wavelength.[9]

There has been some indirect evidence[4,10,11] which may be considered to indicate that interstitial aluminum is sometimes contained in the quartz structure. JOST and ARNOLD found independently that the saturation absorption coefficient of the substitutional aluminum color center system gives poor or no correlation with aluminum content in contrast to earlier agreement discovered by MITCHELL and PAIGE[3] and COHEN[5]. MITCHELL and PAIGE also find enhancement of the substitutional color center after neutron irradiation followed by annealing and X-irradiation.[4] The proposed mechanism for this increase in the substitutional aluminum content is that interstitial silicon and silicon vacancies are produced by neutron bombardment resulting in one interstitial silicon present per silicon vacancy while many times this amount of interstitial aluminum is

* Presented at *Symposium on Defect Structure of Quartz and Glassy Silica*, Mellon Institute, May 20–24, 1957.

† Pittsburgh Plate Glass Company Research Project.

present in the particular crystal originally. Upon annealing, there is competition for the silicon vacancy; and some of the interstitial aluminum becomes substitutional with enhancement of the aluminum band upon X-ray treatment.

BUERGER[12] has discussed the stuffing of the β-quartz structure with β-eucryptite, $LiAlSiO_4$. He compares his X-ray data on β-eucryptite with that of WINKLER[13] who found the space group to be the same as that of β-quartz, namely, $C6_22$, while BUERGER finds P2 (probably). The cell constants of β-quartz are $a = 5{\cdot}02$ and $c = 5{\cdot}48$, while the constants for β-eucryptite are $a = 5{\cdot}27$ (WINKLER) or $2\times5{\cdot}275$ (BUERGER), $c = 2\times5{\cdot}625$ (WINKLER) or $2\times5{\cdot}61$ (BUERGER). In eucryptite the aluminum atoms are substitutional for half the silicon atoms of the β-quartz structure, alternating in layers along the c axis, while an equivalent number of lithium atoms are in the interstitial void tetrahedrally coordinated by oxygen and on level with the aluminum atoms. The interstitial lithium is thus bound electrostatically to the oxygens; and electroneutrality of the structure, containing a Li(I) and an Al(III) in place of a Si(IV), is preserved.

Recent unpublished work of ROY[14] and of VAUGHAN[15] indicates that β-spodumene, $LiAl(SiO_3)_2$, also has the β-quartz structure. It is also possible that further studies of the high temperature forms of petalite, $LiAl(Si_2O_5)_2$, may disclose a third lithium aluminum silicate with the β-quartz structure. In comparing the β-eucryptite and β-quartz structures, one finds that their thermal expansion characteristics are different along the c-axis.[16] The former contracts along the c-axis,[6] while the latter expands.[17] At 650°C the expansion of β-eucryptite is about 0·5 per cent along the a-axis and the contraction is about 1·1 per cent along the c-axis, while at the same temperature β-quartz expands 1·02 per cent along the c-axis and 1·76 per cent along the a-axis.

EXPERIMENTAL

The X-ray irradiations were performed in a Picker Industrial X-ray Unit designed for continuous operation at 50 kVp, 40 mA. The tube used was a Machlett AEG50T with tungsten target and beryllium window. All X-irradiations were made near room temperature. The quartz wafers undergoing transition to the β-phase (one to two mm in thickness) were heated in a nichrome element electric muffle furnace. The samples were cooled slowly by shutting down the furnace and allowing it to cool to room temperature overnight before removing. In some cases the quartz samples were not cracked, in others, one or two cracks would develop.

RESULTS AND DISCUSSION

During an investigation of the aluminum color center system in synthetic quartz a specimen was obtained from L. A. THOMAS of General Electric Co., Wembley, England, grown on a Z-seed plate (General Electric Co. No. 515B). This quartz did not visibly color, even after several billion roentgens of X-irradiation, except in regions of rhombohedral growth. The Z-growth (perpendicular to the c-axis) region, exhibiting no coloration, was analyzed by an emission method giving results shown in Table 1 under the heading "total crystal" and found to contain substantial amounts of aluminum. The aluminum content differed in various portions of the crystal sampled, indicating the aluminum content was not uniform as may be seen comparing columns headed (1) and (2) in Table 1. In one case (2) it was found that the lithium content closely matched the aluminum content in the Z-growth region. Since the smoky colored regions are indicative of substitutional aluminum (and interstitial lithium or sodium) and since the Z-growth developed no color, yet contained considerable aluminum and alkali, it was concluded

Table 1. *Emission analysis, Z-seed, General Electric Co. synthetic quartz* (515*B*) *impurity content per million silicon atoms*

Impurity	Zone		Total crystal*	
	Clear	Precipitated	(1)	(2)
Al	90	135	180	290
Ca	8	11	9	11
Cr	8	6	8	4
Cu	trace	trace	trace	6
Fe	10	11	10	8
Li	26	35	61	260
Na	24	132	106	131

* Excluding regions of rhombohedral growth containing substitutional aluminum.

VOL.
13
1960

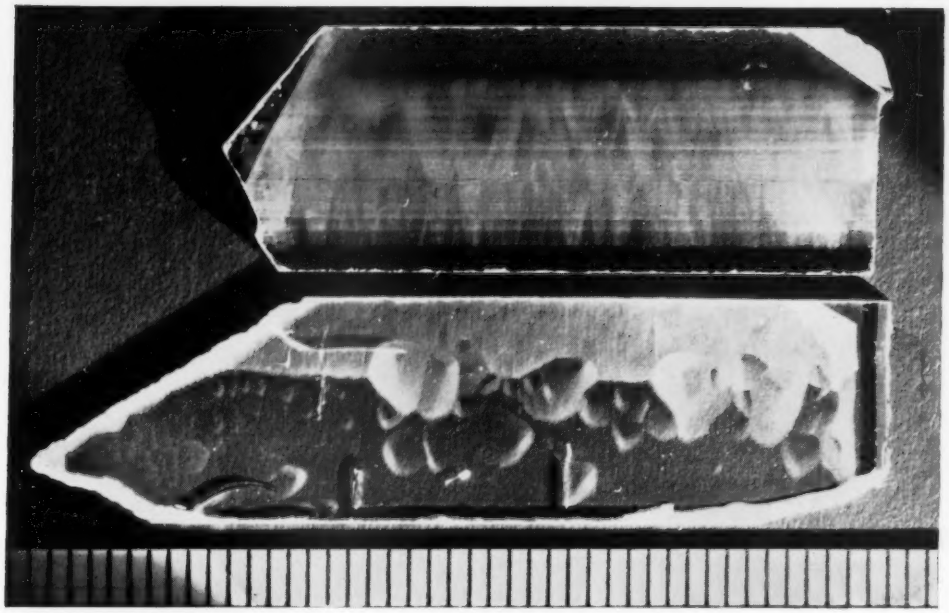


FIG. 1. General Electric Co. synthetic quartz No. 515B grown on *Z*-seed plate after heat treatment to precipitate lithium aluminum silicate. 1 scale division = 1 mm. Lower figure: Wafer cut perpendicular to *c*-axis. Upper figure: Wafer cut parallel to *c*-axis in same portion of crystal as above.

FIG. 2. Photograph of growth figure on rhombohedral face of natural quartz from Herkimer County, New York. This figure is similar to growth figure on 0001 surface of synthetic quartz showing three distinct types of surface.

[*facing p.* 322

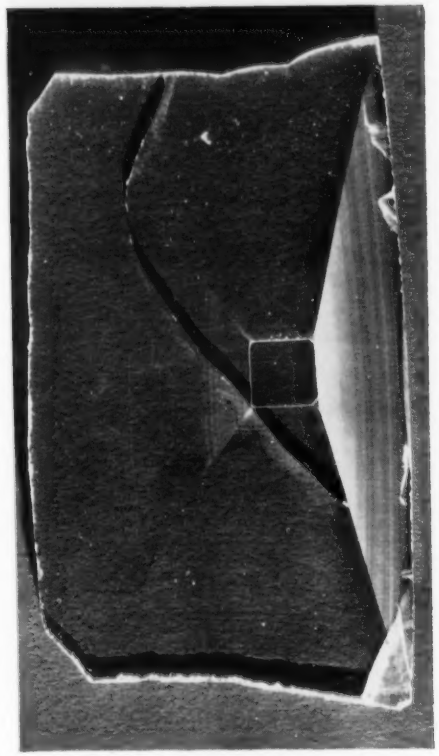


FIG. 3. Wafer of *Y*-bar crystal cut perpendicular to *Y*-axis after precipitation treatment. Major precipitation of lithium aluminum silicate in X_5^- region (see Fig. 5, drawing on right). The crack cutting across illustration is a result of the heat treatment. Magnification the same as in Fig. 6.

FIG. 4. Same wafer as in Fig. 3 with more intense Tyndall effect in order to intensify weaker areas of precipitation barely visible with weaker Tyndall light. Magnification the same as in Fig. 6.

VOL
1
1961

VOL. 13 1960

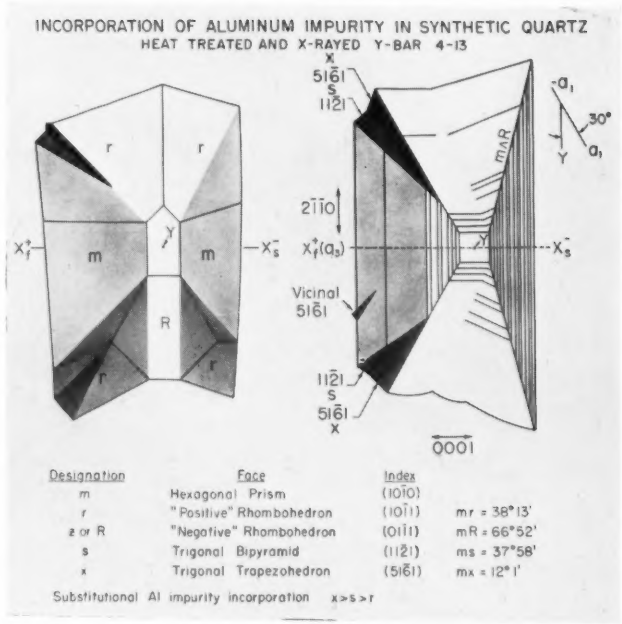


FIG. 5. Artist's representation of substitutional and interstitial aluminum in a *Y*-bar synthetic quartz crystal as elucidated by heat and X-ray treatments.

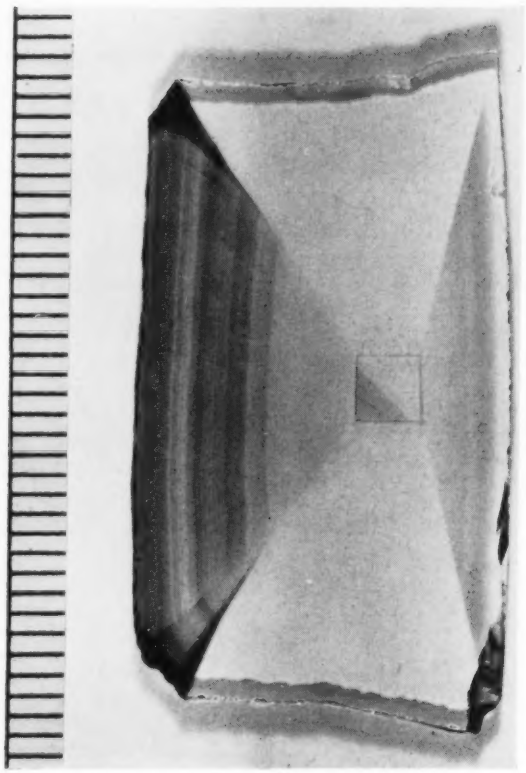


FIG. 6. *Y*-bar from same crystal as in Figs. 3 and 4 which has been X-rayed rather than heated. Visible coloration is due to color center system related to substitutional aluminum impurity. 1 scale division = 1 mm.

VOL. 13 1960

that it was most probable that the aluminum in this region of the crystal is *interstitial.*

At this juncture the author recalled the work of BUERGER[12] on the stuffing of β-quartz by eucryptite and devised an experiment to test the possibility that quartz contains interstitial aluminum. This experiment consisted in cutting and polishing thin wafers of the quartz grown on a *Z*-seed both parallel and perpendicular to the *c*-axis of the crystal. These wafers were slowly heated in an electric furnace through the α–β transition temperature of quartz to 650°C. The wafers were allowed to remain at this temperature for $2\frac{1}{2}$ hr and then cooled slowly down through the α–β transition temperature. The crystal wafers were then allowed to continue to cool slowly to room temperature overnight. The result is illustrated in Fig. 1. A precipitate was produced by the heat treatment in both wafers. Additional experiments showed that the precipitate appears immediately on cooling to the α–β transition temperature of the quartz. The Tyndall effect is used in order to see the fine precipitate produced. The individual particles can just be distinguished at 80× magnification. No precipitation takes place in the portions of rhombohedral growth seen in the upper right corner of the lower wafer and the upper left of the upper wafer. These are the areas that develop the color center system characteristic of substitutional aluminum on X-ray treatment. The precipitate shows clearly the mode of growth of the synthetic quartz on a *Z*-seed plate. Macromosiacs several millimeters in diameter grow from the seed plate (as seen in the upper wafer cut parallel to the *c*-axis) and result in the spiral growth figures on the 0001 surface of the quartz crystal. This 0001 surface can be reproduced in the interior by the precipitation as illustrated in the lower wafer (cut perpendicular to the *c*-axis). These spiral growth patterns have a distorted hexagonal shape and show a threefold difference in amount of precipitation in some cases. This corresponds with the threefold surface orientation of the growth figure as illustrated in Fig. 2.

The upper wafer (Fig. 1) shows that the material that is precipitated is laid down layerwise during growth in different concentrations and for some unknown reason is not incorporated at all into this particular crystal during early growth or during late growth. This phenomenon undoubtedly is related to the aluminum (and possibly lithium) concentration in the feed solution from which the crystal is growing and possibly the temperature of the autoclave.

That this precipitate incorporates cations other than Si(IV) is proven by the increase in the spin–lattice relaxation time of the Si^{29} nucleus in the quartz after precipitation.[18] This indicates that traces of paramagnetic impurity such as chromium and iron are being coprecipitated and are no longer incorporated interstitially in the α-quartz lattice. These paramagnetic impurities, when present in the lattice and in the vicinity of Si^{29} ions, cause the latter to relax faster in a magnetic field. These impurities undoubtedly coprecipitate in Al(III) positions.

Emission analyses of the clear and precipitated zones (Table 1) indicate that the zone where precipitation has taken place contains a greater amount of aluminum, lithium and sodium. This may indicate that the precipitate is a mixed sodium–lithium aluminum silicate. Why the clear zone which contains aluminum, lithium and sodium, does not contain precipitate is not clearly understood. However, it is to be noted that in the clear zone the aluminum content is greater than the sum of the lithium and sodium contents, while in the precipitated zone the alkali contents are slightly greater than the aluminum content. Nor is it clearly understood from where in the lattice the silicon ions come if it is indeed a silicate that precipitates. Attempts to determine the nature of the precipitate by electron diffraction methods have thus far failed.[19]

The precipitation technique was extended to *Y*-bar quartz, obtained through the courtesy of D. HALE of the Clevite Research Center, Cleveland, Ohio, as illustrated in Figs. 3 and 4. Fig. 4 was photographed using a more powerful light beam than in Fig. 3 in order to intensify the weaker lines shown in the former.

It is seen that the major portion of the precipitation takes place in the X_s^- zone (see Fig. 5 for drawing listing various zones) of the *Y*-bar. Table 2 gives emission analyses of the *Y*-bar crystal studied (No. 4–13). It is found that the X_s^- and X_f^+ zones contain the same amounts of aluminum and lithium; however, the X_s^- zone is higher in both iron and sodium than the X_f^+ zone. The crystal is usually grown with the X_s^- side facing up

Table 2. Emission analysis, Y-bar, Clevite synthetic quartz (4–13) impurity content per million silicon atoms

Impurity	Zone			
	X_s^-	X_f^+	Z_1	Z_2
Al	225	225	< 2	24
Cr	< 4	< 4	< 4	< 4
Fe	32	11	32	11
Li	26	26	trace?	trace?
Na	80	trace?	trace	trace?

and in many cases particles of acmite (aegirite), $NaFe(SiO_3)_2$, are incorporated into this zone. The bright white random specks seen scattered in Fig. 3 are of this mineral composition. There are several crystallites clustered along the top of the square seed plate in this figure. They appear as black specks in Fig. 4. In Table 2 it is seen that the *Z*-zones are low in aluminum and alkali. However, aluminum is present and one is able to precipitate a slight amount in these zones as is brought out more clearly in Fig. 4.

Fig. 6 is a photograph of a second wafer from *Y*-bar 4–13 which was X-rayed in order to develop the visible color center system related to substitutional aluminum. It is seen that no color is developed in the *Z*-zones and that aluminum in these zones is therefore all interstitial. It is further seen on comparing Figs. 3 and 4 with Fig. 6 that the areas in X_f^+ that do not color on X-ray treatment are just those areas where precipitation of interstitial aluminum occurs. Why interstitial and substitutional aluminum do not occur in the same regions, or conversely why aluminum is sometimes incorporated interstitially and sometimes substitutionally in quartz, is not entirely clear. The *Z*-zones may be largely free of aluminum because the voids paralleling the *c*-axis allow freer exchange of impurity with the solution during growth. In examining the seed in Fig. 6, one sees that it contains some substitutional aluminum; however, the aluminum incorporation into the growing crystal is independent of that contained in the seed.

MITCHELL and BASTIN[20] have independently found cloudiness induced in General Electric Co. quartz grown on a *Z*-seed plate upon heat treatment. They found that the intensity of the light scattering was dependent on aluminum concentration, temperature and duration of the heat treatment. In this laboratory increasing the duration of the heat treatment was found to increase the spin–lattice relaxation time of Si^{29} indicating that more paramagnetic impurity would coprecipitate with aluminum upon continued heat treatment at 650°C in agreement with the findings of these workers.

In examining Fig. 6 in detail, one observes that the intensity of the substitutional aluminum coloration is not uniform. Two effects are observed:

(1) Many lines in the X_f^+ $(2\bar{1}\bar{1}0)$ region. These are caused by alternating pickup of substitutional aluminum during growth and extend into the *s* $(11\bar{2}1)$ and *x* $(51\bar{6}1)$ regions also. This in many ways resembles the Liesegang ring phenomenon and may be due to variation in concentration of aluminum in the feed solution and to the variation of temperature of the autoclave during growth.

(2) Substitutional aluminum occluded in the growing quartz preferentially in zones in the order $\text{——}\ 51\bar{6}1 > 11\bar{2}1 > 2\bar{1}\bar{1}0 > 0001$.

These zones may be easily located by studying Fig. 5 which is a schematic representation of Figs. 3, 4 and 6. Normally, aluminum does not enter *Z*-growth (0001) substitutionally if only trace amounts are present in the feed solution. The nonuniform distribution of aluminum in these zones is undoubtedly related to its difference in charge from silicon (IV). The electrical charge on the growing face may also be a contributing factor to the order of concentration of impurity incorporation.

Germanium has been found to enter the quartz lattice in a statistical manner[21] as one would predict for an impurity of the same charge as silicon and only slightly larger in ionic size. One unexpected feature in doping a *Y*-bar with germanium is that the incorporation of substitutional aluminum in the *Z*-zone (0001) is promoted, although it still enters the lattice in smaller amounts in this zone than in the other zones.[21] In flawed quartz grown on a *Z*-seed one also finds large amounts of substitutional aluminum because in this case the growth is all rhombohedral, which also accounts for the flawed growth.

In examining the X_f^+ region of Fig. 6 (near top of figure) one observes a few dark streaks radiating out from the body of the crystal (see Fig. 5 for

representation labeled vicinal growth) to the surface. These streaks have the same depth of color as $51\bar{6}1$ or $11\bar{2}1$ growth and the same orientation as one of the two. These radiating lines are very pronounced when observed on a *Y*-bar crystal wafer under the microscope and terminate in the vicinal growth of the $2\bar{1}\bar{1}0$ face. When one measures the orientation of this vicinal growth, one finds that these faces are all *x*- or *s*-faces. Thus, by X-raying the crystal and creating the visible color centers related to substitutional aluminum, one has solved the origin of these vicinal faces which are seen to originate at flaws or dislocations in the $2\bar{1}\bar{1}0$ zone and continue out radially to the surface. Thus, it would be possible to grind the rough X_f^+ surface smooth and then reproduce its rough fine-structure by careful examination of X-rayed *Y*-bar cross-section cuts at intervals along the crystal. The vicinal growth is not related to nor does it start at the seed of the synthetic crystal.

VOL. 13 1960

CONCLUSIONS

Aluminum may be incorporated into α-quartz either interstitially or substitutionally. In either case it is always accompanied by alkali ion, lithium and/or sodium. Substitutional aluminum is incorporated in various growing zones of the quartz in different concentrations generally being absent or interstitial in 0001 growth unless germanium is present, in which case it can be incorporated in the 0001 growth substitutionally. The order of the concentration of alkali follows that of the substitutional aluminum in the various zones.

Interstitial aluminum may be precipitated by heat-treating the quartz in the β-quartz stability range and then cooling. It was found that the precipitate comes out at the α–β transition temperature and is most likely a lithium aluminum silicate phase isomorphous with β-quartz.

Visible color center phenomenon in quartz is related to substitutional aluminum impurity but not to interstitial aluminum.

Acknowledgements—Mr. L. A. THOMAS of General Electric Company Ltd., Wembley, England, is thanked for the *Z*-seed quartz studied and Dr. D. HALE of the Clevite Research Center, Cleveland, Ohio, is thanked for the *Y*-bar quartz used in this work. Mr. M. MASER took the photomicrograph used in Fig. 2 and Mrs. B. S. MILAN performed the emission analyses presented in this paper.

REFERENCES

1. GRIFFITHS J. H. E., OWEN J. and WARD I. M., *Nature, Lond.* **173,** 439 (1954).
2. O'BRIEN M. C. M., *Proc. Roy. Soc.* **A231**, 404 (1955).
3. MITCHELL E. W. J. and PAIGE E. G. S., *Phil. Mag.* **46** (ser. 7) 1353 (1955).
4. MITCHELL E. W. J. and PAIGE E. G. S., *Phil. Mag.* **1** (ser. 8) 1085 (1956).
5. COHEN A. J., *J. Chem. Phys.* **25**, 908 (1956).
6. PAIGE E. G. S., *Phil. Mag.* **2** (ser. 8) 864 (1957).
7. COHEN A. J., *J. Chem. Phys.* **28**, 401 (1958).
8. COHEN A. J., *Glastechn. Ber.* **32K,** Part VI, 53 (1959).
9. COHEN A. J., *J. Chem. Phys.* **23**, 589 (1955).
10. JOST J. M., *Studies of Optimum Methods of Quartz Synthesis*, U.S. Signal Corps Project 142B, 1 March, 1955–29 February, 1956.
11. ARNOLD G. W. JR., *Proc. 10th Annual Symp. on Frequency Control, Asbury Park, N.J., May*, 1956, p. 60.
12. BUERGER M. J., *Am. Min.* **39**, 600 (1954).
13. WINKLER H. G. F., *Acta Cryst. Camb.* **1**, 27 (1948).
14. ROY R., personal communication.
15. VAUGHAN P. A., personal communication.
16. GILLERY F. H. and BUSH E. A., *J. Amer. Ceram. Soc.* **42**, 175 (1959).
17. *Handbook of Physical Constants*, Geological Society of America, Special Papers No. 36, January, 1942, p. 35.
18. HOLZMAN G. R., personal communication.
19. ROY R., personal communication.
20. MITCHELL E. W. J., personal communication.
21. COHEN A. J. and HODGE E. S., *J. Phys. Chem. Solids* **7**, 361 (1958).

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 13. pp. 326–336. Printed in Great Britain.

INFRARED ABSORPTION OF DEFECTS IN QUARTZ*

D. L. WOOD

Bell Telephone Laboratories, Inc., Murray Hill, N.J.

(*Received* 27 *May* 1959)

Abstract—A number of sharp absorption bands in the infrared spectrum of crystalline α-quartz between 3200 cm^{-1} and 3655 cm^{-1} have been assigned to defects in the lattice. Defects containing water may cause a broad absorption through this region of the spectrum in very imperfect parts of certain crystals, but it does not contribute to the sharp bands observed. One band at 3581 cm^{-1} is probably due to the OH stretching vibration of a proton defect, but the other peaks are not. Treatments of the crystal involving X-irradiation, heat bleaching, annealing or heating in an electric field strongly affect the intensities of the bands observed, and it is concluded that the bands originate in color centers with an electronic origin. Probably more than one type of center is involved. No correlation with the visible color centers of smoky quartz or amethyst is observed. Interstitial foreign atoms play an important role in the formation of the infrared centers but vacancies or interstitial Si or O atoms are probably not involved. Although much information relating to the infrared color centers is available the full details of their structure are not yet known.

INTRODUCTION

THE optical absorption spectrum of the pure quartz lattice is characterized by a very broad region of high transparency, beginning with the electronic absorption band which produces the vacuum ultraviolet cutoff near 69,000 cm^{-1} and continuing to about 3000 cm^{-1} in the near infrared region as shown in Fig. 1. This wide region of transparency contributes to the frequent use of quartz for optical elements. From 3000 cm^{-1} toward longer wavelengths there is strong absorption starting with the overtone and combination frequencies from 2500 cm^{-1} to about 1500 cm^{-1} passing through the fundamental lattice vibrations (not all of which are shown in Fig. 1) and ending in the vicinity of 120 cm^{-1} (80μ) with the last strong fundamental lattice absorption band. Beyond 120 cm^{-1} the crystal is again quite transparent. These properties are all characteristic of the pure crystal lattice, and therefore the same in every crystal.

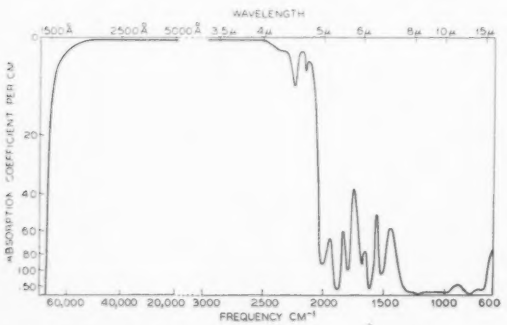


FIG. 1. Optical absorption of quartz.

In addition to the consistent features of the optical spectrum, a very high percentage of crystals exhibit optical absorptions of a variable nature in the otherwise transparent regions. These absorptions are of two types in general, one type being due to the inclusion in the lattice of a chemical species which imparts a color characteristic of the species, as in the case of iron, giving the variety of yellow quartz called citrine. The second type of absorption is usually called a color center and may, for example, give rise to the familiar brownish coloration of smoky quartz or to the violet coloration of amethyst. To the defects causing these two common varieties of colored quartz have been added an ever increasing list of color centers in the visible and ultraviolet regions of the spectrum. Some of these color centers are understood in considerable detail while others, including those in the

* Paper presented at *Symposium on Defect Structure of Quartz and Glassy Silica*, Mellon Institute, May 22–24, 1957.

infrared, remain as yet unexplained. Those defects causing the visible and ultraviolet color centers are not the primary concern of this article and will only be discussed as they bear on the primary problem of the origin of the infrared absorbing centers. The question of the origin of the infrared bands is important in the investigation of the differences which apparently exist between natural quartz and synthetically produced quartz[1] and which must originate in defects in the crystal lattice.

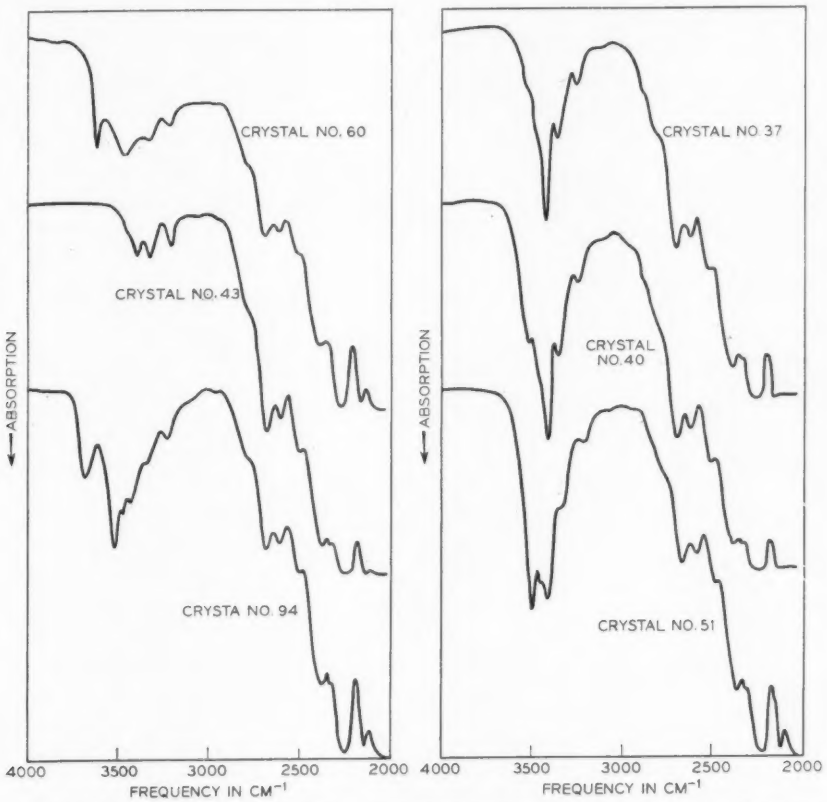


FIG. 2. Infrared bands in quartz at room temperature. NaCl prism, basal sections.

GENERAL FEATURES OF THE INFRARED BANDS NEAR 3400 cm^{-1}

It has been known for a long time that the absorption coefficients of a group of bands near 3400 cm^{-1} in the infrared spectrum of quartz were not the same in every crystal.[2,3] It has recently been recognized[4,5] that the reason for the apparent failure of Lambert's law arises from the fact that the intensity of the bands arises from a defect in the crystal lattice rather than from a proper vibration. It is therefore not correct to assign these bands to overtone or combination bands as several authors[6–8] have done.

A survey of nearly 100 quartz crystals of various origins has revealed several facts about the appearance of the near infrared bands in question. In the first place, there are several components to the group, and the intensities of the several components occur with varying magnitude. Six common characteristic contours of the group of bands for basal sections of various crystals are shown in Fig. 2 as recorded at room temperature using a

Perkin–Elmer double-beam Model 21 spectrophotometer with NaCl optics. The crystals numbered 37, 40, 43 and 51 are clear Brazilian quartz selected from commercial radio-grade material. The crystal numbered 60 is synthetic quartz grown by the Bell Telephone Laboratories on what has been called a *Y*-bar seed. The sample numbered 94 is from a very fine Brazilian rose colored natural crystal courtesy of Dr. A. J. COHEN of the Mellon Institute, Pittsburgh, Pennsylvania.

The detailed picture of the individual components and their relative intensities is more clearly revealed by the spectra at 77°K which are shown on an expanded scale in Fig. 3 for the same crystal sections whose room temperature spectra are shown in Fig. 2. In these spectra recorded with the Perkin–Elmer Model 21 spectrophotometer with a CaF_2 prism are distinguishable at least twelve individual lines whose frequencies are listed in Table 1. In the body of the table the

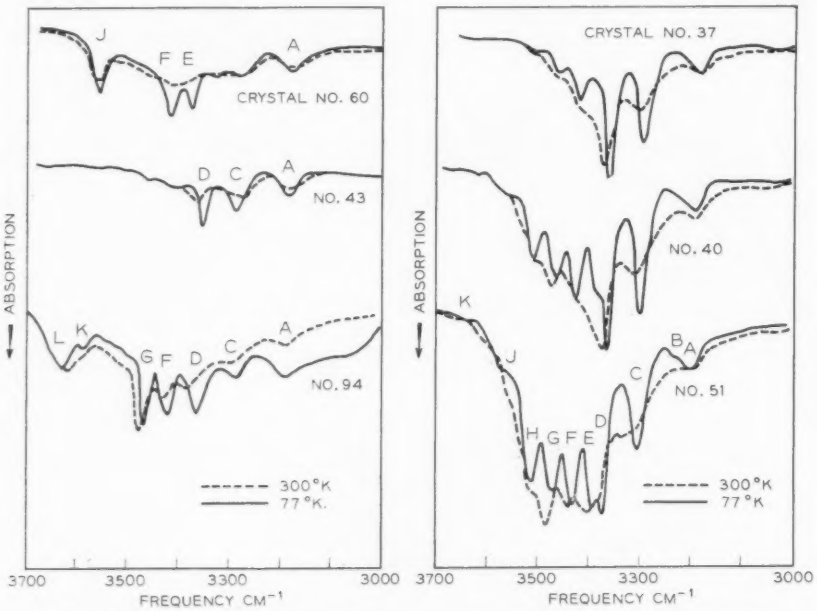


FIG. 3. Infrared bands in quartz at low temperature. CaF_2 prism, basal sections. Dashed curve: room temperature. Full curve: 77°K.

The absorption bands in Fig. 2 which vary in intensity are found at frequencies higher than 2700 cm^{-1}. The several absorption peaks at lower frequency when reduced to absorption coefficients are quite constant and can be assigned to proper overtones and combinations of the fundamentals which lie below 1250 cm^{-1}. It is immediately obvious from Fig. 2 that there are many individual components in the group of variable bands and not all components are found with the same intensity in the various crystals. The intensity of the strongest band in the group ranges from 0·13 cm^{-1} to 6·0 cm^{-1} in the crystals which have been examined so far.

presence of a given line is designated by a + opposite the crystal number, and the lack of a line is designated by no symbol. The designations of the lines given in this work are noted across the top of the table together with the designations given by MITCHELL and RIGDEN[4] for comparison. Most quartz crystals give spectra resembling those of crystals numbered 37, 40, 43 and 51, although the details of the pattern may vary somewhat. The over-all aspect at 77°K is rather similar for these crystals even though the room temperature spectra may appear quite different. All synthetic crystals and the rose quartz, however, seem to have characteristics not in common with those of the natural

Table 1. *Absorption lines in quartz at* 77°K

Line designation												
This work	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*	*K*	*L*
MITCHELL and RIGDEN	ε		δ		γ	β	α			χ		
Frequency (cm^{-1})	3200	3234	3305	3370	3393	3432	3470	3512	3563	3581	3634	3654
Crystal Nos.												
37	+	+	+	+		+	+	+				
40	+	+	+	+	+	+	+	+	+	+		
43	+		+	+		+						
51	+	+	+	+	+	+	+	+		+		
60	+		+	+	+	+				+		
94	+		+	+		+	+			+	+	+

crystals just mentioned. The spectrum from crystal number 60 shown both in Fig. 2 and in Fig. 3 indicates that the high frequency band designated *J* has considerable intensity, while the band designated *C*, which is the dominant feature for common varieties of natural quartz is very weak in the synthetic quartz spectrum. The spectrum of the rose quartz is unique since no other natural or synthetic crystal has a spectrum even roughly comparable. Because of the variability of intensities of these individual components it does not appear logical to assign all the lines to a common origin, and one may conclude that the bands originate from more than one defect.

Even though the spectra at 77°K show considerably more detail than the room temperature spectra there is a great deal of detail yet to be resolved as the temperature is lowered. BASS and BROIDA[9] have shown that as the temperature is decreased to 4·2°K the width of the individual components may decrease to less than 2 cm^{-1}. At this temperature they have identified two lines in crystalline quartz (*D* and *G*) which we also find to be present in the spectra of our crystals. The increased detail in the spectrum when the sample temperature is lowered to that of liquid helium and when LiF is used as a prism material instead of CaF_2 is shown in Fig. 4 where the spectrum of a basal section of crystal number 51 is given. Seven of the absorption bands in the spectrum of this crystal at 77°K as shown in Fig. 3 are now split in such a way as to give sixteen in Fig. 4. The limit of the detail in the spectrum of Fig. 4 is due to the resolving power of the spectrometer employed and it appears that more components would be visible if the higher resolution of a diffraction grating were used.

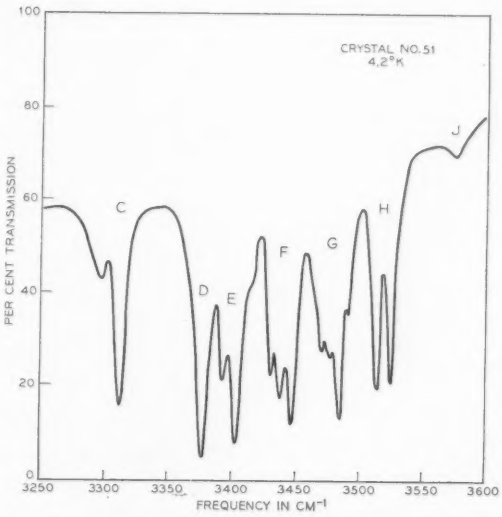


FIG. 4. Infrared bands in quartz at liquid helium temperature. LiF prism, basal section.

The variability of the infrared absorption bands is not only limited to the distinction between one crystal and another, but is also found within a single crystal. In most natural crystals the variation in the infrared bands from one point to another is

a matter of intensity alone, perhaps closely related to the segregation of impurities which causes the smoky color phantoms often seen. In synthetic crystals the variation of the infrared bands may involve the general contour of the bands as well as their intensity. This is perhaps not surprising in view of the fact that the smoky color center introduced in synthetic crystals on X-irradiation are also

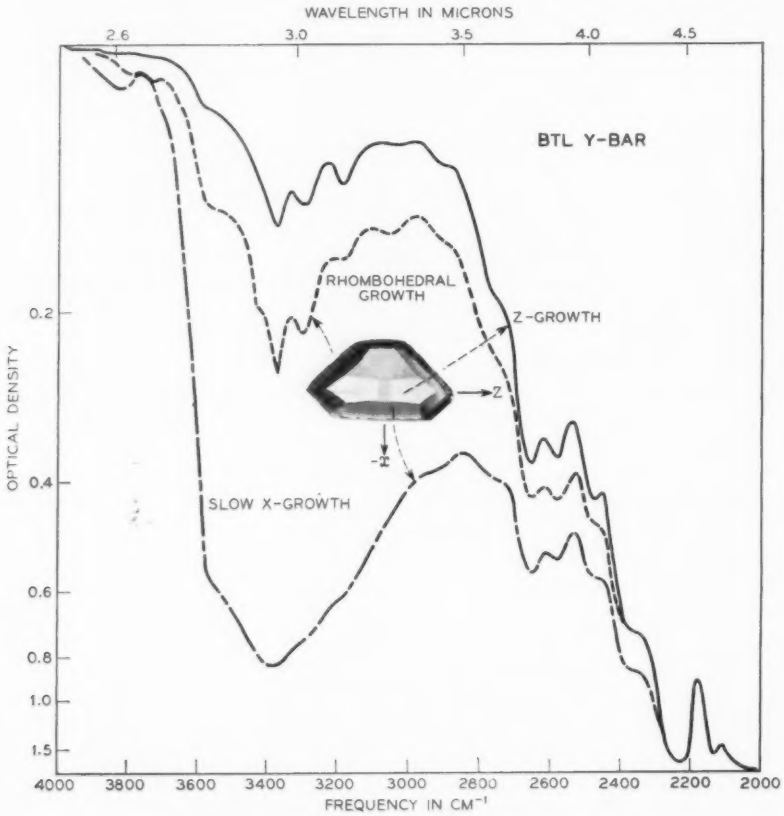


FIG. 5. Segregation of defects in quartz. Section $\perp$ y-axis.

segregated and evidently of more than one kind. In the center of Fig. 5 is shown a photograph of an X-irradiated section of a synthetic crystal cut perpendicular to the Y-axis. The rhombohedral growth sections are very dark, but the color centers are concentrated near the surface which was next to the X-ray source. The fast and slow X-growth parts of the crystal are lighter in color, but in these portions the color centers are distributed throughout the section uniformly. In the part of the crystal which grew along the z-axis there are no color centers at all. The small rectangle in the center of the photo is the cross section of the seed on which the crystal was grown. The infrared spectra of three of these parts of the crystal are indicated in Fig. 5 by the three curves identified by the arrows with the appropriate part. The part of the crystal which grew from the seed along the z-axis has the smallest absorption for the frequencies in question, while that part which grew on the rhombohedral face has intermediate intensity. The part which grew along the negative X-direction (slow x-growth) has a very broad diffuse spectrum. This segregation of the defects responsible for the spectroscopic inhomogeneities is an important factor which must be considered in experiments

on all quartz crystals, but especially for the synthetic material.

INFRARED BANDS AND LATTICE DEFECTS

Several types of defect in the lattice could be considered in attempting to account for the infrared absorption bands. We will consider five possibilities in some detail: (i) Water molecules may be included in voids in the lattice to produce the observed absorption bands; (ii) a substitutional impurity or (iii) an interstitial impurity may give rise to either a molecular or an electronic transition of this energy; (iv) protons may be bound to the lattice to form OH groups whose stretching vibration would give the observed frequency and dichroism;[4] (v) a vacancy in the lattice may have around it an electronic system within which a transition of this very low energy may take place.

The possibility of water in the quartz lattice actually was the starting point for the investigation described here. It was suggested that this defect might be found in greater concentration in synthetic quartz, and thus be the cause of the difference in electromechanical properties between synthetic and natural crystals. We now believe that the sharp lines observed are not due to water in the lattice, but it is possible that the broad diffuse absorption of the slow x-growth material shown in Fig. 5 is due to water inclusions. The contour of the bands are similar in milky quartz where the myriad of inclusions causing the scattering of light and the milky appearance may contain liquid water. Similarly in synthetic quartz the slow x-growth material contains a very high concentration of inclusions which may also contain water.[10] The other regions of the crystal do not contain these inclusions and they also do not have the broad diffuse infrared absorption. Further data is necessary, however, before this conclusion can be accepted unequivocally. In particular, studies of synthetic crystals grown in D_2O should be useful.

There is a possibility that the infrared bands might be due to overtones and combinations of fundamentals of molecular units such as LiO or NaO which could be formed from impurities in covalent bonding with the oxygen of the lattice. This is however untenable since the spectra of various crystals do not differ in any other region of the infrared spectrum except the 3400 cm^{-1} region. If such molecular units were producing these overtones, they would also necessarily produce fundamentals at lower frequency, and these are not observed. Furthermore, the level of impurities in the quartz lattice is in the range from 1 to 100 ppm, and the intensity of any overtone or combination frequency involving these impurities would be several orders of magnitude weaker than the overtones and combination of the pure lattice which are found nearby (e.g. at 2660 cm^{-1}). The bands actually observed are more intense than the nearest lattice overtones or combinations, and therefore cannot have this origin.

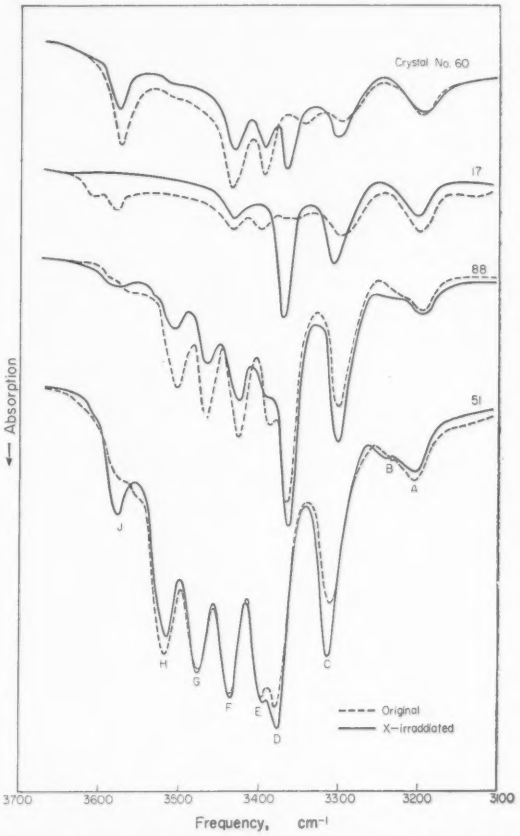


FIG. 6. Effect of X-irradiation on the infrared bands of quartz at 77°K. Basal sections. Dashed curve: before irradiation. Full curve: after irradiation.

It is difficult to assign all the infrared bands of quartz to the proton defect. In the first place there are many components to account for. This would

require the postulation of at least sixteen discrete stable OH-group configurations in the lattice, each with slightly different vibrational frequency. This does not seem likely. Furthermore the experiments to be described later are difficult to explain on this basis. It is possible that one component (*J*) at 3581 cm^{-1} may be due to OH as suggested by MITCHELL and RIGDEN[4]. But because the band is weak, our data do not warrant a firm conclusion on this point.

Strong evidence against the OH assignment comes from experiments in which the intensities of the infrared bands were altered but only electrons in the lattice were affected. In the first of these experiments basal sections of the crystal were subjected to about 10^6 r of X-rays from a tungsten target tube operated at 50 kVp and 50 mA with a beryllium window. The results of this treatment for four crystals are shown in Fig. 6. It is evident that the change of intensity for the various components of the infrared absorption is not the same. Some bands decrease in intensity and others increase in intensity during the treatment. The results are listed in Table 2 where the band designation and the frequency of the band are given together with the symbol designating no change (0), an increase (+) or a decrease (−). The bands of extreme frequency, namely *A*, *B*, *K* and *L* (*K*, *L* not shown in Fig. 6) do not change on X-irradiation. The lines *C* and *D* increase in intensity while the rest decrease in intensity. From the crystals examined so far it appears that the largest change of intensity occurs in the synthetic quartz crystals exemplified by Nos. 17 and 60 rather than in natural crystals.

The effect of X-irradiation on the infrared spectrum can be completely reversed by heating the crystal to 500°C for a few minutes. This is the second experiment effecting electrons in the crystal. Any change in intensity with this treatment which will be called heat bleaching proceeds only to the extent of reversing the changes due to X-irradiation. When an unirradiated crystal is given this mild heat treatment no change in intensity of the lines is observed. This behavior is strongly reminiscent of the visible color centers previously mentioned since they also can be removed by a mild heat bleaching.[11,12] These experiments suggest an electronic origin for the infrared bands since the X-irradiation and the heat bleaching presumably affect only electrons within the crystal.

Since the infrared bands behave in a manner similar to that of the visible color centers in regard to X-irradiation and bleaching it is logical to ask if there is a correlation between the visible color centers and the infrared centers. Since none of the crystals having infrared bands were violet in color, it is concluded that these bands do not bear any relation to the color center of amethyst.

Table 2. Changes in the infrared bands with X-irradiation

Line	Frequency (cm^{-1})	Intensity change on X-irradiation	Line	Frequency (cm^{-1})	Intensity change on X-irradiation
A	3200	0	*G*	3470	−
B	3234	0	*H*	3512	−
C	3305	+	*I*	3563	−
D	3370	+	*J*	3581	−
E	3393	−	*K*	3634	0
F	3432	−	*L*	3654	0

Some crystals, however, were smoky when received and practically all exhibited the smoky color center after X-irradiation. This color center both in natural and X-irradiated smoky crystals has been attributed by GRIFFITHS *et al.*[13] to an electronic transition in the vicinity of a substitutional aluminum atom in the lattice. The maximum absorption of the ordinary ray in smoky quartz has been observed by most workers at 460 mμ wavelengths. In our crystals however, we have observed two different colors in basal sections of smoky quartz either of natural origin or of X-irradiated material. The grey smoky color characterized by the 460 mμ peak is common but there

is also a yellowish variety with an absorption maximum near 400 mμ. The visible absorption spectra of sections of two crystals cut perpendicular to the *c*-axis and exhibiting these absorption bands are shown in Fig. 7. Crystal No. 51 shows the 400 mμ

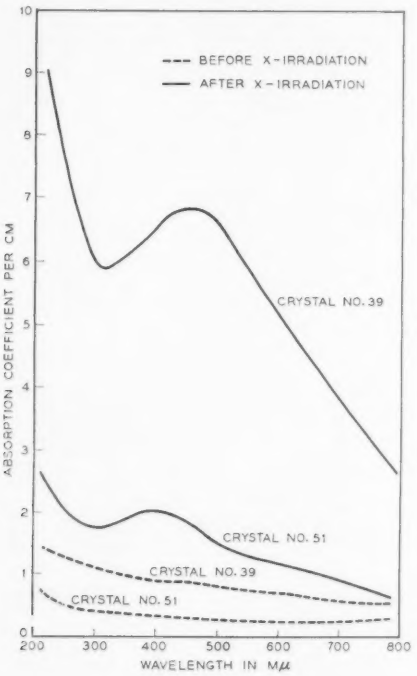


FIG. 7. Absorption curves of two different colors of smoky quartz produced by X-irradiation of two natural Brazilian crystals. Basal sections.

absorption peak, whereas crystal No. 39 shows the 460 mμ peak. Neither of these color centers, however appears to be related to the infrared color centers as can be seen from the data in Table 3. In this table the intensity of the visible smoky color center is given next to the intensity of the principal peak of the infrared color center, namely the band *C* at 3305 cm^{-1}. The two crystals No. 40 and No. 48 show this lack of correlation. Crystal No. 48 is rather dense after X-irradiation but its infrared absorption is weak, whereas crystal No. 40 darkens very little on X-irradiation and yet its infrared absorption is relatively strong. Another pair of crystals which show a similar lack of correlation between the visible color center and the infrared absorption is No. 39 and No. 70. We conclude

Table 3. Intensities of infrared and visible absorption-bands

Crystal Nos.	Intensity after 3 hr *X*-irradiation		Intensity at 3305 cm^{-1} (cm^{-1})
	460 mμ	400 mμ	
17	3·9	0·5	0·30
37			1·00
38	2·8		0·32
39	6·9		0·14
40		0·9	1·65
43	1·0		0·21
45		1·25	1·13
48		1·95	0·50
51		2·0	0·78
70	1·2		3·64

therefore that the infrared absorption is not connected with the visible color centers and is not due to substitutional aluminum in the lattice.

It should be mentioned in passing that in several crystals we have observed dichroism in the basal section for the visible smoky color centers. This is, of course, impossible for a defect having the symmetry of the quartz lattice, and therefore, leads us to the conclusion that the aluminum color center in these crystals is segregated into sub-microscopic domains having symmetry different from that of the crystal as a whole. It is likely that the orientation of the domains is related to the direction of growth of the crystal. This interesting phenomenon deserves further study.

VACANCIES AND INTERSTITIAL ATOMS

MITCHELL and RIGDEN[4] have eliminated oxygen vacancies, silicon vacancies, and the corresponding displaced interstitial atoms as defects responsible for the infrared bands because neutron irradiation decreases the intensity of the bands. Our work has not yet included direct experiments involving these defects, and we accept their conclusions as proved.

Another experiment bearing on the origin of the infrared bands involves heating the crystal to 950°C for several hours, which causes the relative intensity of the various infrared bands to change as shown in Fig. 8. The intensity of the high frequency components is decreased while that of the low frequency components is increased. This is in

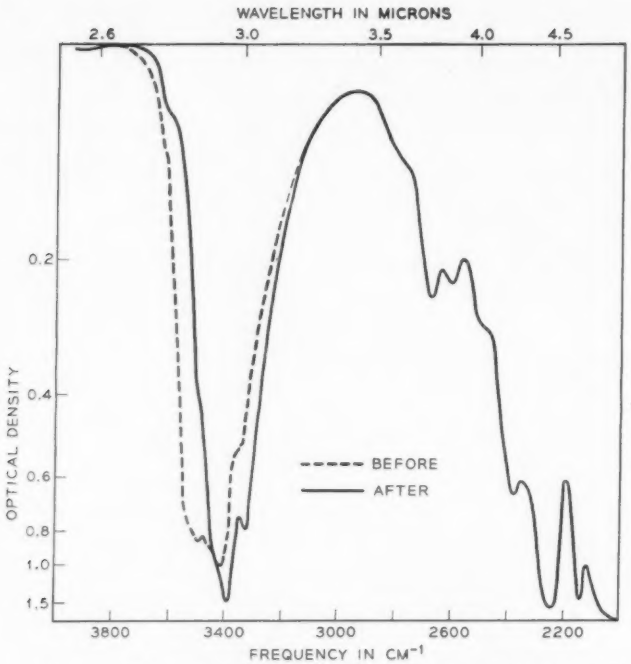


FIG. 8. Effect of annealing at 950°C for 4 hr on the infrared bands of quartz. Basal section.

contrast to the heat treatment at 500°C where no change in the infrared spectrum is observed except in previously X-irradiated specimens. At 950°C the temperature is high enough so that migration of interstitial ions and perhaps the annealing of vacancies may take place, since it has been observed that the intensity of the X-ray color centers in the visible region is increased by the high temperature annealing treatment. This is shown in Fig. 9 where spectra have been recorded from an X-irradiated basal section of crystal No. 51 and from an identical section of the same crystal which had been heated to 950°C for several hours, cooled, and X-irradiated for the same length of time. The intensity of the absorption in the visible for the annealed sample is much greater than for the control which was not annealed. This suggests that aluminum atoms which originally were located in the interstitial voids of the crystal have succeeded in displacing silicon atoms from the lattice, forming substitutional defects. The increase in the number of substitutional aluminum atoms in the lattice, therefore, accounts for the increase in the intensity of absorption of the visible color center. It will be noticed also that the peak of the visible absorption band has been shifted from about 400 mμ to about 460 mμ due to the annealing treatment. This suggests that the 400 mμ center is not due to substitutional aluminum.

A second set of experiments having a direct bearing on the origin of the infrared bands involves heating a crystal to 500°C in an electric field of about 3000 V/cm parallel to the c-axis of the crystal. When this is done, a considerable electric current flows through the crystal due to the diffusion of ions through the lattice voids parallel to the crystallographic c-axis. Many workers have shown[14–16] that the current which flows when the crystal is heated in the field is due to the migration of alkali cations. The current which flows when the crystal is first put into the field starts at a large value, builds up to a maximum, then decreases over a period of minutes to a small fraction of its original value. After continuing to decrease more slowly over a period of hours, a condition is reached where for several days a very small current flows.

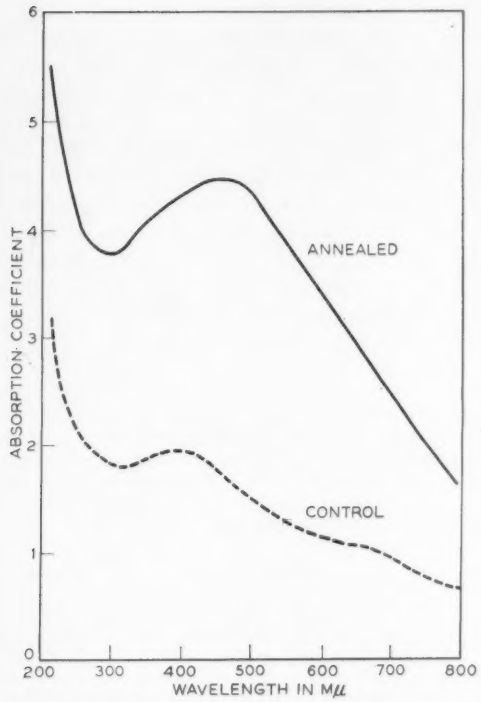


FIG. 9. Shift in frequency and peak intensity of the visible absorption band of quartz after annealing and X-irradiation. Basal sections.

The initial high current value corresponds to an electrical polarization of the crystal, while the second phase corresponds to a flow of alkali cations through the axial voids. It is believed[15] that the long-term small current may be due to electrolysis of the crystal accompanied by evolution of oxygen and deposition of silicon.

The effect of the electro-diffusion on the infrared absorption band is appreciable after an extremely short treatment. In Fig. 10 are shown the spectra taken from four samples of crystal No. 70, each sample having a different duration of electro-diffusion treatment. After 2 min at 500°C in a field of 3000 V/cm, the intensities of the various components are appreciably altered, the most noticeable effect being the broadening of the bands, *G* and *F* lying in the vicinity of 3470 cm^{-1} and 3515 cm^{-1}. After 45 min of electro-diffusion, bands marked *G* and *F* have maxima appreciably shifted from the original values due to the change of intensity of the components of these bands. The bands marked *A*, *B* and *D* have not changed intensity while those marked *E* and *H* have decreased in intensity until after 9 hr of treatment *H* is completely missing. The band marked *C* on the other hand has increased somewhat in intensity.

Since the principal effect of the electro-diffusion treatment is the removal of interstitial ions from the lattice, it is logical to conclude that the interstitial ions play an important role in the origin of the infrared absorption bands. Spectrographic analyses of the diffused and undiffused areas of a crystal show that all the lithium and most of the sodium have been removed from the crystal during electro-diffusion treatment. This is in agreement with the findings of ADAMS[14] and of WENDEN[15] who found that the diffusion rate of lithium was extremely high and that of sodium was somewhat lower.

The interstitial alkali cation also plays an important role in the visible color centers of quartz, since the intensity of the visible absorption bands introduced by X-irradiation is negligible if the crystal has first been treated by electro-diffusion.

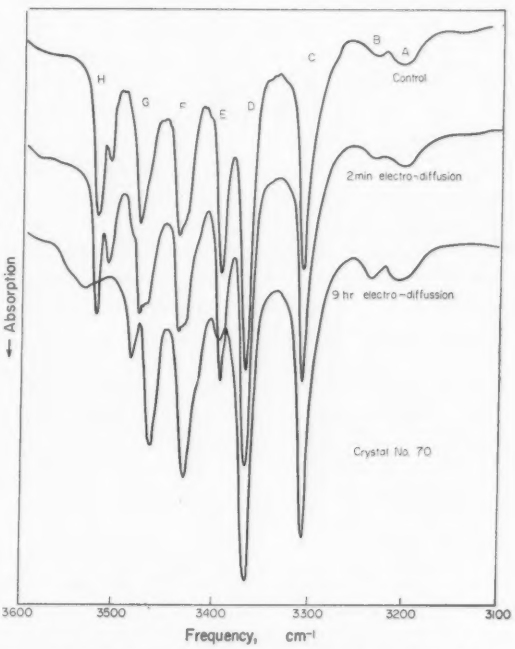


FIG. 10. Effect of electro-diffusion at 500°C on the infrared bands of quartz as recorded at 77°K. Basal sections.

The areas of the crystal which have been under the electrode during the electro-diffusion treatment will not exhibit the smoky color center after X-radiation while the areas which were not under the electrode during the treatment do exhibit the color center on X-radiation.

O'BRIEN[17] has postulated that when the visible color center is formed on X-irradiation of a crystal having substitutional aluminum in the lattice, the electron removed from the aluminum atom by the incident photon is taken up by the interstitial sodium or lithium ion which balances the charge of the substitutional aluminum atom before irradiation. If this is true, the removal of interstitial cations by electro-diffusion before X-irradiation may mean that the photo-electron can no longer escape from the aluminum center, and the color center is not formed. The observation that electro-diffused crystals do not darken on X-irradiation, therefore, lends support to O'BRIEN's model.

CONCLUSIONS

It is not possible at the present time to determine in detail the nature of the defects causing the 3 μ infrared bands in quartz. It seems certain, however, that the bands are electronic in origin since the X-irradiation and heat bleaching experiments affect the infrared intensities, but only electrons are involved in the treatment of the crystal. From several possible types of defect which might be considered, MITCHELL and RIGDEN[4] have eliminated interstitial displaced oxygen and silicon and the corresponding vacancies because the infrared bands decrease in intensity on neutron bombardment. Foreign atoms in the lattice must therefore be responsible for the bands. From the possible foreign atom species we have ruled out substitutional aluminum, because of the lack of correlation between the intensities of the aluminum color center in the visible and the infrared bands. The annealing and electro-diffusion experiments emphasize the importance of interstitial alkali cations, principally sodium and lithium. These are important in the formation of the visible color centers as well as the infrared centers, as the same experiments show. It seems unlikely that interstitial protons can account for the infrared bands even though the OH stretching frequency is the same magnitude as that of the observed bands. We believe that in the case of crystals containing appreciable concentrations of liquid inclusions, the water in the inclusions may contribute a broad diffuse envelope superimposed on the sharper defect absorption bands. The final details of the centers causing the infrared bands will become apparent only after further experiments.

Acknowledgements—I am greatly indebted to Miss D. M. DODD for a large part of the recording and processing of spectrophotometric data, to R. A. LAUDISE for many fruitful discussions and for a collection of synthetic quartz crystals, and to Dr. A. J. COHEN for the gift of a section of Brazilian rose quartz.

REFERENCES

1. KING J. C., *Phys. Rev.* **109**, 1552 (1958).
2. DRUMMOND D. G., *Proc. Roy. Soc.* **A153**, 328 (1936).
3. DREISCH T., *Z. Phys.* **42**, 426 (1927).
4. MITCHELL E. W. J. and RIGDEN J. D., *Phil. Mag.* (ser. 8) **2**, 941 (1957).
5. WOOD D. L., *J. Chem. Phys.* **27**, 1438 (1957).
6. PLYLER E. K., *Phys. Rev.* **33**, 48 (1929).
7. ARNOLD G. W., *J. Chem. Phys.* **22**, 1259 (1954).
8. SCHAEFER C. and MATOSSI F., *Das Ultrarote Spektrum.* Springer, Berlin (1930).
9. BASS A. M. and BROIDA H. P., *J. Opt. Soc. Amer.* **47**, 163 (1957).
10. BROWN C. S., KELL R. C., THOMAS L. A., WOOSTER N. and WOOSTER W. A., *Min. Mag.* **29**, 858 (1952).
11. MITCHELL E. W. J. and PAIGE E. G. S., *Phil. Mag.* (ser. 8) **1**, 1085 (1956).
12. COHEN A. J., *Amer. Min.* **41**, 874 (1956).
13. GRIFFITHS J. H. E., OWEN J. and WARD I. M., *Rep. Bristol Conf. on Defects in Crystalline Solids* p. 81. Physical Society, London (1955).
14. ADAMS J. A. S., Ph.D. Thesis, University of Chicago (1951).
15. WENDEN H. E., *Amer. Min.* **42**, 859 (1957).
16. VERHOOGEN J., *Amer. Min.* **37**, 637 (1952).
17. O'BRIEN M. C. M., *Proc. Roy. Soc.* **A231**, 404 (1955).

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 13. pp. 337–343. Printed in Great Britain.

THE EFFECT OF IMPURITIES ON THE GROWTH OF SYNTHETIC QUARTZ*

C. S. BROWN and L. A. THOMAS

Research Laboratories, The General Electric Company, Wembley, England

(*Received* 27 *May* 1959)

Abstract—An investigation has been carried out aimed at developing a process for the growth of large crystals of synthetic quartz using silica-rich rocks native to the United Kingdom as the raw material. As a result of this work the effect of impurities on the habit and quality of the crystals has been determined. The role of aluminium as a key impurity has been proved and the darkening produced by X-irradiation has been used to study the distribution of impurities in the synthetic quartz. Recent work has thrown further light on the manner in which aluminium, possibly associated with another element, is incorporated in the quartz structure.

VOL. 13 1960

INTRODUCTION

UNTIL recently, all the quartz required for the production of oscillator crystals for frequency control has been obtained from natural resources. Although quartz is one of the most abundant minerals in the earth's crust, it is only found in large crystals of the required quality in a few isolated regions. During the past 15 years work has been carried out to develop processes for the controlled growth of quartz in the laboratory. Considerable success has been obtained both in the United Kingdom[1,2] and the United States,[3,4] crystals of piezoelectric-quality weighing over 1 lb. having been grown.

All the modern processes used for the growth of synthetic quartz have been developed from the pioneer work of SPEZIA[5] at the turn of the century. Although the methods adopted by various workers in this field are basically similar, there are certain fundamental differences which affect the properties of the crystals. Because of its glass-forming properties and its allotropic modifications it is not possible to grow quartz from the vapour or the melt. Growth from solution must be used and as quartz is virtually insoluble in aqueous media under ordinary ambient conditions it is necessary to use elevated temperatures and pressures to obtain sufficient solubility. These so-called hydrothermal conditions are probably similar to those in which much natural quartz has been formed. At temperatures approaching 400°C and pressures of 1000 atm (about 7 tons/in.^{2}), quartz is readily soluble in alkaline solutions such as sodium carbonate.

CRYSTAL GROWING TECHNIQUE

A schematic diagram of the apparatus used by the authors is shown in Fig. 1. In what will be called the standard process a steel autoclave constructed to withstand high pressures has seed crystals suspended from the lid and a supply of crushed melting-grade quartz at the bottom. The autoclave is about 80 per cent filled with a solution containing 88 g/l of sodium carbonate and sealed. The simple furnace used consists of a hotplate on which the autoclave stands surrounded by micaceous-flake thermal insulation. By this means a temperature gradient is established so that it is hotter at the bottom in the region of the nutrient crushed quartz than at the top where the seeds are located.

At the working conditions the autoclave will be filled with a single-phase fluid, the pressure developed being a function of the temperature and of the percentage of the space originally occupied by the solution at room temperature. The temperature at the base of the autoclave is controlled

* Paper presented at *Symposium on Defect Structure of Quartz and Glassy Silica*, Mellon Institute, May 22–24, 1957.

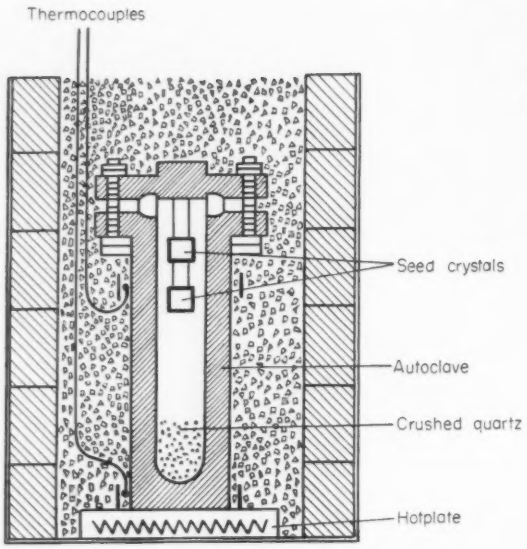


FIG. 1. Diagram of apparatus for the growth of synthetic quartz.

at about 400°C and the temperature at the seeds reaches an equilibrium some 40°C lower. The temperature gradient along the length of the autoclave is not uniform, a fall of about 20°C occurring across the metal at the base and most of the remaining drop being across the nutrient. The space above the nutrient is approximately isothermal and the supersaturation in this region remains constant. Thus crystals can be grown at approximately the same rate of growth in any part of the autoclave.

For a given design of autoclave, the rate of growth will be dependent on seed orientation, pressure, temperature and temperature differential. The seed orientation which has been adopted in most of the work to be described is the basal plane or *Z*-cut. Fig. 2 shows the relation between this cut and the minor rhombohedral or *r*-cut which has also been used as a seed for the growth of synthetic quartz, particularly in the United States. It will be shown subsequently that the seed orientation not only affects the rate of growth but also has a marked effect on the way in which impurities are incorporated in the crystal. It is most convenient to control the rate of growth by means of the temperature difference and this is done by adjusting the flow of solution through the nutrient either by grading the particles of the nutrient itself or by the use of an additional baffle above the nutrient. In order to obtain crystals of the highest visual quality, the rate of growth must not exceed a certain maximum, which, for the conditions used in the standard process, is about 0·5 mm/day on each side of the seed measured in the direction of the optic axis. In fact the visual quality is in some respects a misleading criterion and the measurements of the mechanical damping recorded in Table 1, show that a progressive improvement in crystalline perfection takes place as the growth rate is reduced. As it happens, for most practical applications the quality corresponding to 0·5 mm/day is adequate but for especially stringent requirements it may be necessary to employ a

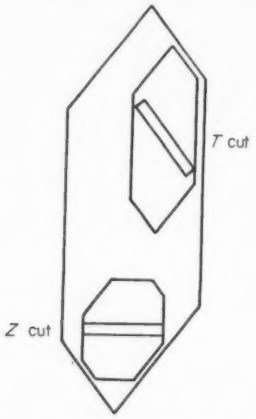


FIG. 2. Relationship between crystals grown on the basal plane (*Z*-cut) and the minor rhombohedral plane (*r*-cut).

Table 1. *Mechanical damping of quartz at* 5 Mc/s

Material	Growth rate (mm/day)	Damping* ($= Q^{-1}$)
		$\times 10^{-7}$
Natural	?	3–4
Synthetic	0·55	30
	0·38	11–15
	0·25	5

* The damping was measured at 20°C using a special design of oscillator crystal.[12]

lower growth rate or an alternative growing technique.

A small pilot plant has been set up to grow crystals by the standard process. A growth of 15 mm takes place in a period of about a month, the resulting crystals weighing about 135 g.

INFLUENCE OF IMPURITIES

The standard process recrystallizes a low-grade quartz—which due to size and imperfections such as twinning is unsuitable for piezoelectric use—into crystals of a size and quality which are ideally suited to this purpose. However, the melting-grade quartz used as nutrient still has to be imported and considerable work has been carried out over the last 5 years to develop processes which can be used with relatively impure nutrient materials which are native to the United Kingdom. Reasonable success has been obtained using flint and a variety of quartzites.

Early in this investigation it was found that quartz could be grown on *Z*-cut seeds using impure nutrient materials, but that the quality, particularly of large crystals, was not good enough for piezoelectric use. It is now known that the poor quality is due to the incorporation of impurities in the synthetic crystal during growth. However, it was found that it was possible to grow crystals of piezoelectric quality by modifying the solution in which the crystals were grown. Good crystals have been grown from flint and impure quartzite by using a solution containing 40 g/l. Na_2CO_3, 33 g/l. NaOH and 3·4 g/l. NaF.

Further examination of a number of different quartzites showed that many of them could be used satisfactorily with the standard process; these are referred to as class *A* quartzites. It is now known that the difference between class *A* and class *B* quartzites (class *B* being those quartzites which require the modified process) lies in the type of accessory minerals which are associated with the quartz. In particular, the structure of the felspar, which is commonly the primary accessory mineral in quartzites, plays an important role in deciding whether the material will be class *A* or *B*.[6]

Work by GRIFFITHS *et al.*[7] using a paramagnetic resonance technique has shown that aluminium is the impurity which affects the quality of synthetic quartz crystals grown from class *B* materials using the standard sodium carbonate solution. This result has been corroborated by spectrographic analysis on a number of specimens. In addition, the direct test has been made by deliberately adding aluminium in a number of forms along with a pure quartz nutrient. Such adulterated quartz now acts as a class *B* nutrient and crystals grown from it using the standard solution have the habit and poor quality of a crystal grown under similar conditions from flint or a class *B* quartzite. Further, by using the modified solution, the defects can again be overcome. Fig. 3 shows four crystals which illustrate this result. The crystals from left to right were grown from nutrients and in solutions as given below.

Nutrient	*Solution*
Quartz	Sodium carbonate
Flint	Sodium carbonate
Quartz+aluminium	Sodium carbonate
Quartz+aluminium	Sodium carbonate+sodium hydroxide+sodium fluoride

The most notable feature which can be seen from the photograph is the nature of the growth surface approximating to (0001). This surface is smooth on the crystals grown from quartz in sodium carbonate solution and from quartz with aluminium in the modified solution, but it is rough and pitted in the case of the crystals grown from flint and from quartz with aluminium, in sodium carbonate solution. It will be shown that the nature of the growth surface is closely related to the manner in which aluminium is incorporated in synthetic quartz grown on *Z*-cut seeds.

It has been known for some time that when natural quartz is irradiated with X-rays or any other ionizing radiation, the material darkens. FRONDEL studied this phenomenon in detail and showed that the darkening produced in natural quartz is often non-uniform showing a banded structure.[8] The darkening occurs in sheets parallel to the major rhombohedral planes and is therefore connected with the growth of the crystal, probably being associated with changes in the environment in which the crystal grew. When large crystals of synthetic quartz were first grown by the authors, their behaviour under X-irradiation was determined. Fig. 4 shows the result of irradiating an X-cut section of a synthetic quartz crystal grown on a *Z*-cut seed. It will be seen that the central

VOL. 13 1960

region corresponding to the natural quartz seed has darkened uniformly and that there are two regions beneath the minor rhombohedral faces which have also darkened rather more intensely than has the seed. The remainder of the synthetic growth has not darkened under this dose (ca.10^7 r). The diagram in Fig. 5 shows the region under the minor rhombohedral face on a larger scale. The triangular region *abc* corresponds to growth which has taken place on the minor rhombohedral face as it develops. Spectrographic analyses of material taken from various regions of a number of crystals grown by the standard process have shown that the total aluminium concentration in the growth on the *Z*-cut orientation is commonly less than 40 parts in 10^6 atomic replacement. However, in the growth under the minor rhombohedral face the aluminium content may be 10 times this figure, i.e. 400 parts in 10^6. The impurity content of the melting-grade nutrient lies between these two figures and the low aluminium concentration in quartz grown on a *Z*-cut seed is in part due to the "scavenger action" of the growth on the minor rhombohedral faces which are formed during the growth. This also illustrates a general result found in the growth of synthetic quartz, namely that it is easier to introduce impurities during growth on the minor rhombohedral face than on the basal plane.

If, however, there is a large excess of aluminium in the system, it is found that the crystals grown on the basal plane will darken readily in a characteristic way as shown in Fig. 6. In addition to the darkening of the seed crystal and the growth under the minor rhombohedral face, there is darkening of the primary growth in the form of distorted narrow-angled cones directed along the *c*-axis. These cones terminate in the rough, pitted growth surface and are apparently associated with the pits. This can also be seen from the bands which run parallel to the seed surface. These bands which are regions of either more or less intense darkening than the surrounding material are reproduced precisely on both sides of the seed. It can, therefore, be deduced that these bands are produced either by changes of the temperature or pressure in the autoclave or, what is more likely, by changes in the concentration of aluminium in the solution. These bands will therefore represent the nature of the growth surface at the particular time when they are formed. The discontinuities in these bands tend to follow the boundaries between the cones of darkening.

MECHANISM OF GROWTH

The characteristic cone darkening structure can be used to obtain a model for the mechanism of growth on the basal plane. This is illustrated in Fig. 7. It is assumed that growth takes place independently on a large number of centres in contrast to growth on a habit face where only a limited number of centres are active and growth takes place by sheets spreading across the growth surface. The growth centres on the basal plane may be associated with spiral dislocations but there is, as yet, no direct evidence for growth spirals. If it is now assumed that quartz is deposited on the individual centres at different rates it will be seen that those which are growing faster overtake their more slowly growing neighbours and render them inactive. In the schematic diagram in Fig. 7 all nine centres are active at the beginning of the growth, at a later stage only centres 2, 4, 6 and 8 are active and eventually only centres 4 and 6. If the aluminium incorporation is uniform on any one growth centre but differs from one to the next, it will be seen that this gives rise to the characteristic darkening pattern. Fig. 7 also shows photomicrographs of the growth surfaces at various stages in the growth together with darkening patterns at similar stages for sections cut perpendicular to the *c*-axis. The individual "cobbles", which are the termination of the growth cones in the surface of the crystal, and the fine structure in the darkening patterns both tend to become coarser as the growth proceeds. This is in agreement with the suggested growth mechanism.

The nature of the rough, pitted growth surface obtained when quartz is grown on the basal plane in the presence of a large excess of aluminium cannot be directly explained in terms of the suggested mechanism. Fig. 8a is a photomicrograph of the surface and shows clearly that the pits have no obvious crystallographic orientation. However, crystals have been grown with only just sufficient aluminium present in the system to commence the incorporation of aluminium by the above mechanism. In this case, the aluminium apparently only goes into the growth on isolated centres and produces a growth surface of the type shown in Fig.

VOL.
13
1960

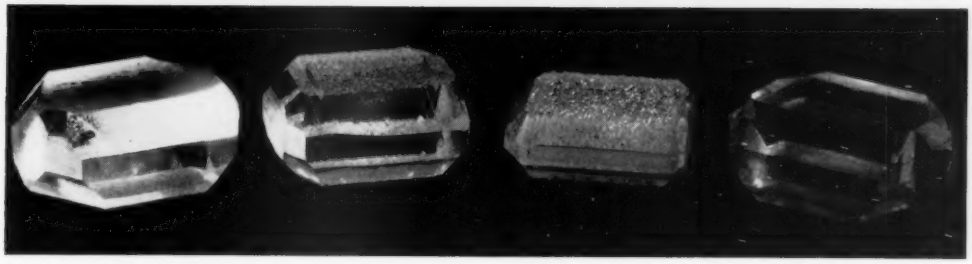


FIG. 3. Crystals grown to illustrate the importance of aluminium as an impurity in low-quality nutrients.

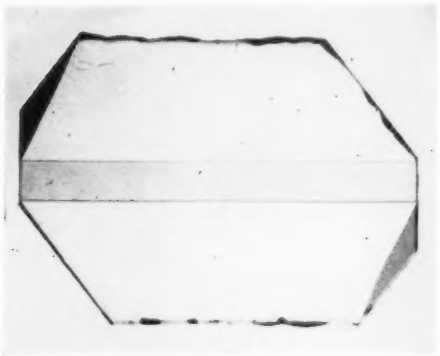


FIG. 4. Photomicrograph of an *X*-cut section from a synthetic quartz crystal grown on a *Z*-cut seed from a pure melting-grade quartz nutrient after *X*-irradiation.

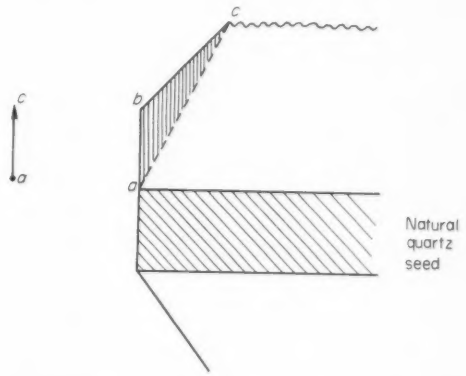


FIG. 5. Diagram showing the X-ray darkening of quartz deposited in the accessory growth on the minor rhombohedral face.

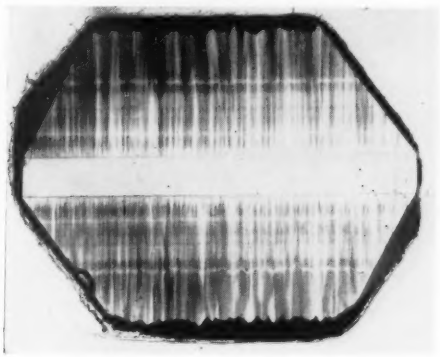


FIG. 6. Characteristic X-ray darkening pattern found in an *X*-cut section of a synthetic quartz crystal grown on a *Z*-cut seed in the presence of a large excess of aluminium.

[*facing p.* 340

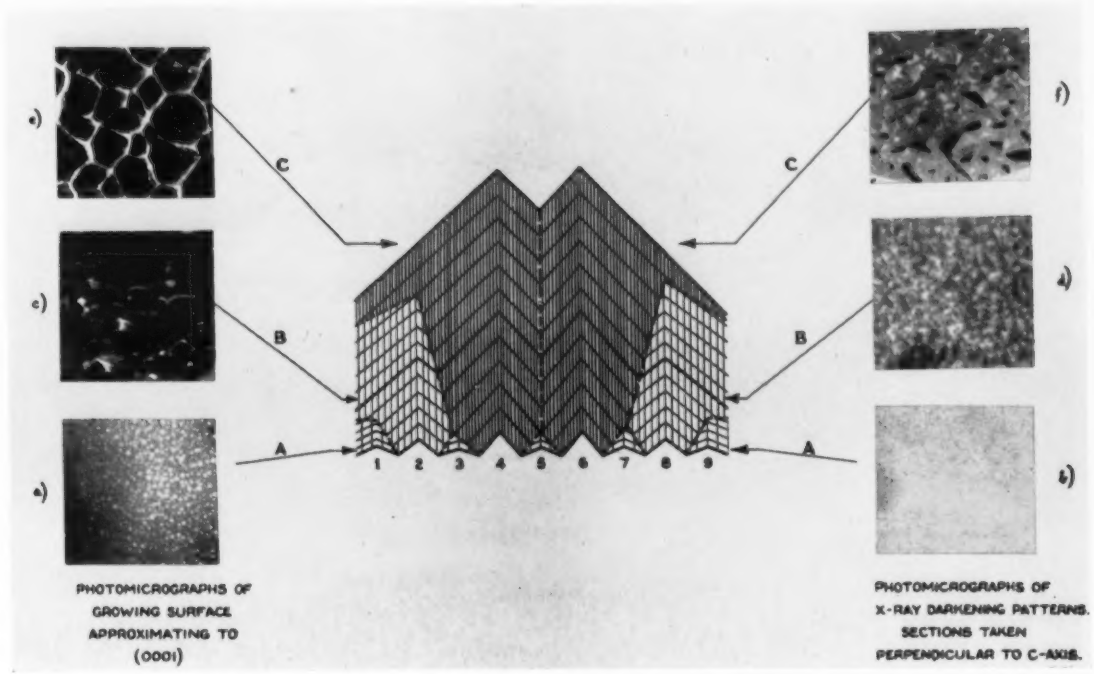


FIG. 7. Illustrating the mechanism of growth of synthetic quartz on *Z*-cut seeds.

VOL.
13
1960

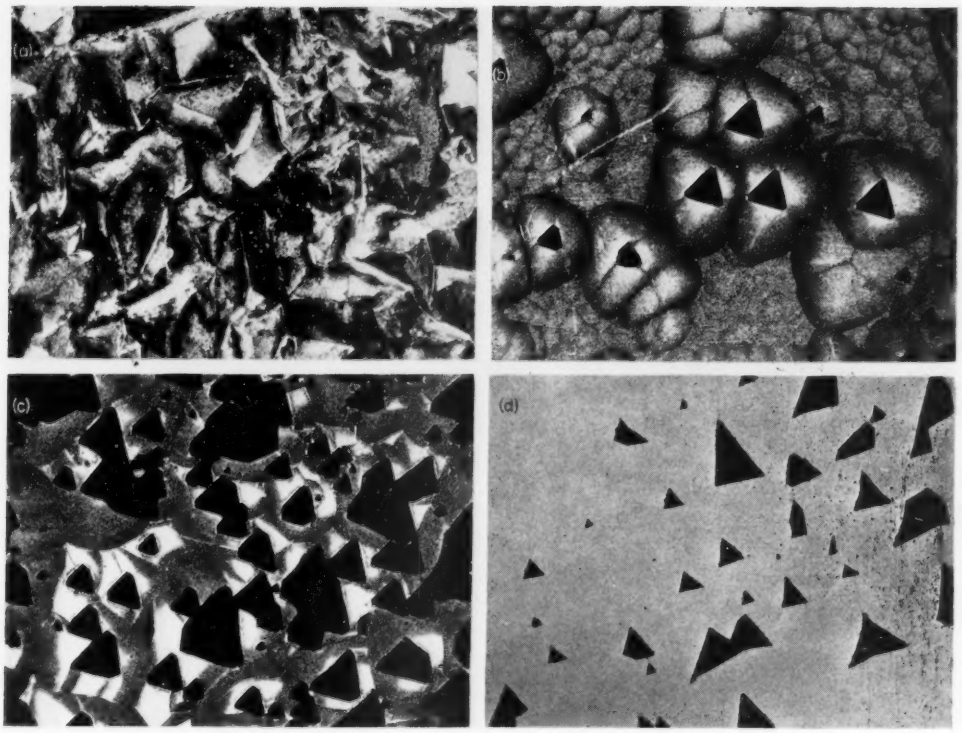


FIG. 8. The nature of the rough growth surface formed in the presence of a large excess of aluminium. (a) Rough growth surface. (b) Isolated pits. (c) Interfering pits. (d) Rough surface after most of the damage has been removed by lapping.

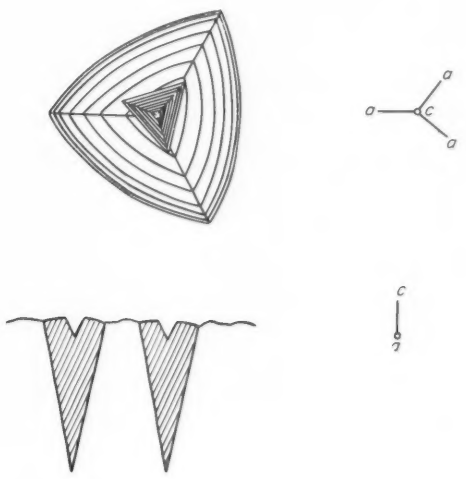


FIG. 9. Diagram of a pit formed at the centre of a growth cobble.

FIG. 10. Photomicrograph of an *X*-cut section showing the relationship between the cones of darkening and the pits.

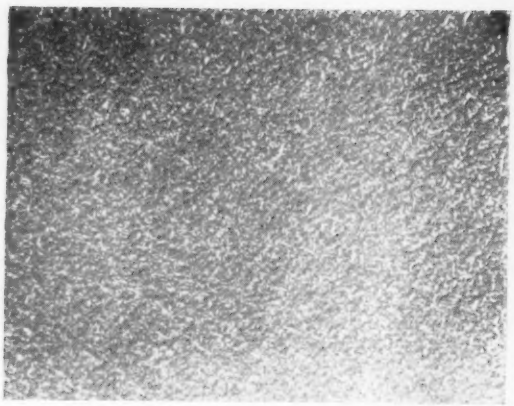


FIG. 11. Strain pattern observed in crossed nicols in a *Z*-cut section of a crystal containing a large excess of aluminium. Section taken near seed growth interface.

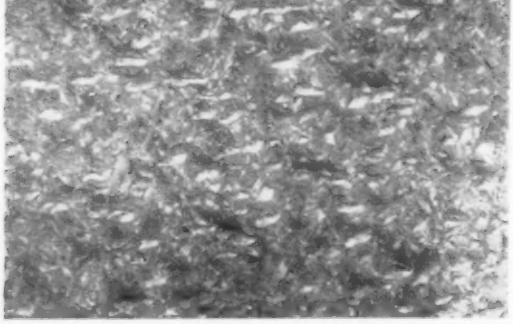


FIG. 12. As Fig. 11. Section taken near end of growth.

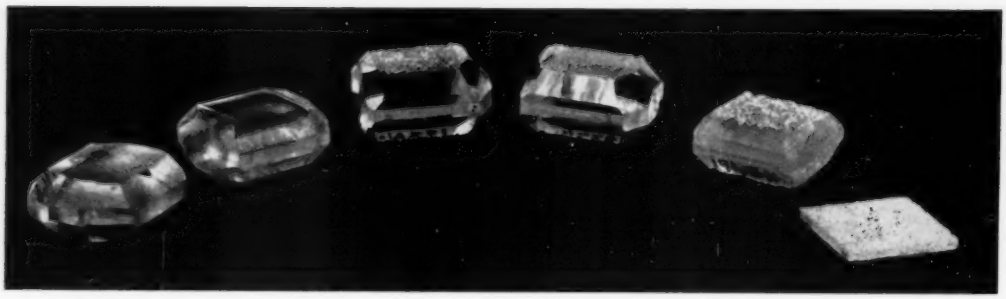


FIG. 13. Crystals grown in the presence of increasing concentrations of aluminium. Aluminium by weight of quartz nutrient: (1) 0·05 per cent, (2) 0·125 per cent, (3) 0·25 per cent, (4) 0·50 per cent, (5) 1·25 per cent, (6) 2·50 per cent.

VOL. 13 1960

8b. It will be seen that certain of the larger cobbles have triangular pits at their centre. The relation between these pits and the rough growth surface produced by a large excess of aluminium is shown in Fig. 8c and 8d. Fig. 8c shows the growth surface of a crystal grown in the presence of a somewhat larger concentration of aluminium than that of Fig. 8b. The pits are here more numerous and are commencing to overlap so that they interfere and lose their obvious crystallographic orientation. The crystal shown in Fig. 8d has a rough surface similar to that in Fig. 8a and has been lapped to remove most of the disturbed surface. The bottoms of the pits of the rough surface show approximately the same form as those in Fig. 8b. The characteristic rough surface is, therefore, formed from a large number of pits which overlap and interfere until the shape and symmetry of the pits is completely lost.

The nature of the individual pits is more readily seen by reference to Fig. 9, which shows a diagram of a single cobble and its pit. The pit shows perfect trigonal symmetry, the sense of rotation from the diad axes being dependent on the hand of the quartz. The sides of the pit form reasonably flat faces which have been indexed using a microscope. It is found that the orientation of these faces corresponds to the indices $25\bar{7}7$. The reason for the formation of such a high index face, if it is a true face, in the presence of excess aluminium is not understood. Fig. 9 also shows a section containing the *c*-axis. If as appears likely, the pit is associated with the incorporation of aluminium, on irradiating such a section the growth cones giving rise to the cobbles with pits at their centres will darken readily. This is shown diagrammatically in Fig. 9 while Fig. 10 shows a photomicrograph which clearly illustrates this phenomenon.

It has been seen that when crystals are grown in a large excess of aluminium, this aluminium is not incorporated uniformly in the crystal as it grows but is taken up preferentially on certain centres at the expense of the neighbouring centres. As the lattice spacing will be a function of the aluminium concentration it is reasonable to expect that strains will be set up at the boundaries between regions of different aluminium content. This strain can be seen readily by examining sections cut either parallel to or perpendicular to the *c*-axis in a polarizing microscope between crossed nicols.

Parallel to the *c*-axis there are deep fissures found under the rough growth surface where the stress exceeds that necessary to produce fracture. Instead of the crystal appearing dark as it should in the extinction position, the field is crossed by bands of light and dark produced by the strain. The same phenomenon is, perhaps, more readily studied by examining sections cut perpendicular to the *c*-axis. Depending on the angle between the polarizer and analyzer the section should appear to be a uniform colour when viewed in white light. Instead, a section cut from a crystal which has non-uniform aluminium incorporation will have a mottled appearance. This is shown in Fig. 11 and 12, which are photomicrographs of two sections of the same crystal, Fig. 11 being taken near the seed crystal and Fig. 12 near the end of the growth. This again illustrates the way in which the number of active growth centres decreases during the growth. In this particular example, the density of active centres has decreased by a factor of the order of 20 in about 5 mm of growth along the *c*-axis.

Comparison of the results of experiments in which crystals are grown on *Z*-cut seeds in the presence of high and low concentrations of aluminium shows that the mechanism by which aluminium is incorporated in the growing crystal is dependent on the concentration of impurity in the system. This has been studied more closely by a set of controlled experiments in which the concentration of aluminium, added as γ-alumina, was steadily increased. Six crystals grown in the presence of aluminium deliberately added to the nutrient in proportions varying from 0·05 to 2·50 per cent by weight are shown in Fig. 13. The concentration of aluminium added to the nutrient for the six crystals numbered 1 to 6 from left to right is given in the caption. Examining the nature of the growth surface it will be seen that the first two crystals show no obvious signs of the inclusion of aluminium. By contrast crystals 3–5 show the characteristic rough growth surface, crystal 5 being so strained that the growth is hardly single-crystalline. With 2·5 per cent aluminium added to the nutrient (crystal 6) all growth is prevented. Further experiments carried out using concentrations of aluminium in the range 0·1–0·25 per cent show that the results are not consistent, in that a number of experiments carried out with the same

aluminium concentration sometimes give a rough growth surface corresponding to non-uniform impurity incorporation and at other times give the smooth cobbled surface of the pure crystal.

This behaviour can be explained in terms of the suggested growth mechanism as follows. For low concentrations of aluminium in the system, the incorporation apparently takes place uniformly. Most of this aluminium is probably interstitial as the material only darkens slightly under X-irradiation. For high concentrations, it has been seen that the aluminium is taken up preferentially on certain growth centres. This can be understood, when it is considered that the energy required to introduce an impurity atom is a function of the number of impurity atoms already incorporated in the growth on this centre. As the number of impurity atoms incorporated increases, the distortion of the lattice becomes greater and it becomes easier to include more impurity atoms. Thus, once the concentration of aluminium in the growth on a given centre exceeds a certain figure, further aluminium atoms will tend to be taken up preferentially on this centre at the expense of the neighbouring centres. In a physical system of this type, the probability that non-uniform inclusion will take place will be a rapidly changing function of the concentration of impurity in the nutrient. It will be seen that this qualitative analysis explains the observations. For low concentrations of aluminium, the probability of non-uniform take-up is very small. As the concentration is increased the stage is reached where there is a reasonable chance that the aluminium is incorporated non-uniformly. This is the region where the results will not be consistent. At still higher concentrations, non-uniform inclusion will be the rule.

Specimens cut from the crystals shown in Fig. 13 have been analyzed for aluminium and sodium using spectrographic and flame photometric techniques, respectively. The results are given in Table 2 in terms of the percentage atomic replacement of silicon by these elements and show that the concentrations of sodium and aluminium are of the same order for low concentrations of aluminium added to the nutrient but that the sodium content tends to saturate while the aluminium content continues to increase. It seems reasonable to conclude that the substitutionally added aluminium is associated with a sodium atom situated interstially. This centre would be responsible for the visible darkening produced by X-irradiation (MITCHELL and PAIGE's[9] *A* bands), a model consistent with that suggested by O'BRIEN and PRYCE[10]. The excess aluminium found when the concentration added to the nutrient is large, could be present either as interstitial atoms or as two substitutional atoms associated with an oxygen vacancy. There is, at present, no evidence to distinguish between these two alternatives.

Table 2. Sodium and aluminium concentrations in synthetic quartz grown in Na_2CO_3 *solution on Z-cut seeds in the presence of aluminium*

Concentration of aluminium added to nutrient (weight per cent)	Impurity concentrations in crystals (atomic per cent)	
	Sodium	Aluminium
—	0·0050	0·0045
0·05	0·0045	0·0065
0·125	0·017	0·045
0·25	0·037	0·065
0·50	0·039	0·11
1·25	0·039	0·29
2·50	—	—

In discussing the incorporation of impurities in synthetic quartz, some mention must also be made of work which has been carried out in attempts to include impurities other than aluminium. It is well known that small monovalent ions in particular lithium and sodium, can readily be introduced into the quartz lattice under the action of an electric field. These ions lie interstitially in the "tunnels" which are parallel to the *c*-axis in the quartz structure. It has already been shown that sodium is present in all synthetic quartz grown by the authors. Attempts have been made to introduce a number of other elements which might be expected to substitute for silicon in the lattice. In general, it has been found extremely difficult to introduce impurities into quartz grown on the basal plane. This result would appear to be different from that found by other workers in this field who have used seeds cut parallel to the minor rhombohedral face. In addition to aluminium, attempts have been made to incorporate the following elements in growth on *Z*-cut seeds.

Group III	Group IV	Group V
Boron	Titanium Germanium Tin	Phosphorus Arsenic

Of these elements only germanium has been successfully incorporated. Bearing in mind the similar ionic radii of silicon and germanium it is not surprising to find that germanium will readily go into quartz as a substitutional impurity. Large amounts of germanium can be taken up by the quartz lattice without setting up measurable strain. As would be expected, the centre is not sensitive to X-irradiation. It is interesting to note that boron is not taken up, although its small size and its valency of three would, at first sight, make it an ideal atom for incorporation in the quartz lattice. No explanation is known for this behaviour.

A recent investigation, still incomplete, has shown that the nature of the cation in the solution from which the crystals are grown can have a considerable effect on the way in which impurities are incorporated in synthetic quartz. A striking example has been the growth of an intensely coloured emerald green quartz, a variety which does not occur naturally. From the circumstances in which this crystal was grown, it has been deduced that the colouration is produced by the presence of a trace of chromium. This diagnosis has still to be confirmed. The colouration in this crystal is very stable, being unaffected by heat treatment up to the α–β inversion temperature or by prolonged X-irradiation. Preliminary transmission measurements show that the material has an apparent cut-off in the ultra-violet at 2800 Å. It appears likely that this synthetic material is not related to the "greened" amethyst described by COHEN[11].

An investigation of the type described can be of value to the worker studying colour centres in quartz in a number of ways.

(1) The process described for the growth of large crystals of synthetic quartz can provide material with a total impurity content, and particularly a substitutional aluminium content, lower than is found in natural quartz. This is of value to the worker studying radiation damage in quartz. With regard to the substitutional aluminium content it must be noted that the concentration in a number of specimens grown under nominally similar conditions will differ slightly, partly as a result of the statistical nature of the process and also as a result of variations in the purity of the nutrient.

(2) Controlled amounts of the impurities present in natural quartz can, in certain instances, be introduced. This can be of considerable help in identifying the nature of those colour centres which are impurity-dependent.

(3) By the introduction of impurities not found in natural quartz, material with new properties can be grown. For example, if sufficient chromium can be introduced into the green quartz it may be possible to obtain a quartz which is paramagnetic.

Acknowledgements—The authors wish to express their gratitude to Mr. J. W. RYDE, F.R.S., of the Laboratories for his close interest and stimulating encouragement of this work. They are also indebted to Dr. E. W. J. MITCHELL of the University of Reading for valuable discussions, especially in regard to the nature of the colour centres occurring in quartz.

REFERENCES

1. BROWN C. S., KELL R. C., THOMAS L. A., WOOSTER N. and WOOSTER W. A., *Min. Mag.* **29,** 858 (1952).
2. BROWN C. S. and THOMAS L. A., *Proc. Inst. Electr. Engrs.* **104c,** 174 (1957).
3. WALKER A. C. and BUEHLER E., *Industr. Engng. Chem.* **46,** 1670 (1954).
4. HALE D. R., *Brush Strokes,* December 1952.
5. SPEZIA G., *Atti Accad. Torino,* **44,** 95 (1908).
6. BROWN C. S., KELL R. C., MIDDLETON P. and THOMAS L. A., *Nature, Lond.* **175,** 602 (1955).
7. GRIFFITHS J. H. E., OWEN J. and WARD I. M., *Rep. Bristol Conf. on Defects in Crystalline Solids* p. 81. Physical Society, London (1955).
8. FRONDEL C., *Amer. Min.* **30,** 432 (1945).
9. MITCHELL E. W. J. and PAIGE E. G. S., *Proc. Phys. Soc. Lond.* B**67,** 262 (1954).
10. O'BRIEN M. C. M. and PRYCE M. H. L., *Rep. Bristol Conf. on Defects in Crystalline Solids* p. 88. Physical Society, London (1955).
11. COHEN A. J., *Amer. Min.* **41,** 874 (1956).
12. BÖMMEL H. E., MASON W. P. and WARNER A. W., *Phys. Rev.* **102,** 64 (1956).

VOL. 13 1960

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 13. pp. 344–346. Printed in Great Britain.

TOPOGRAPHY AND ETCH PATTERNS OF SYNTHETIC QUARTZ*

F. AUGUSTINE and D. R. HALE

Electronic Research Division, Clevite Corporation, Cleveland 8, Ohio

(*Received* 27 *May* 1959)

Abstract—The surfaces of synthetic quartz crystals show many details of growth patterns. Growth hillocks with spiral ramps are often found on rhombohedral faces and on the irregular surface of growing basal pinacoids (00·1). In some instances small hillocks are replaced through etching by pits in a substantially one-to-one relationship.

In contrast to its inertness under ordinary conditions, quartz may be readily dissolved in alkaline solution at elevated temperature and pressure. Use of this fact has led to practical processes for growing well-formed crystals of high quality and purity up to sizes limited only by the size of the pressure vessel.[1–3]

In a typical growing operation, broken clear quartz chunks are placed at the bottom of a vessel having an inside diameter of 2 in. and an inside length of 24 in.; seed plates cut from high quality crystals are suspended in the upper part of the vessel; the vessel is filled three quarters full with 10 per cent (aqueous) sodium carbonate solution; the vessel is closed and heated from underneath to an average temperature of 350°C and raised to a pressure of 15,000 psi. These conditions may be maintained for a period of several weeks. Quartz dissolves at the bottom and the solution is carried upward by convection, where, through cooling, the solution becomes supersaturated. Growth rates may be controlled through a number of variables. Furthermore, seed plates can be cut to present selected crystal planes to the nutrient solution.

The study of spiral growth patterns on various crystals and their interpretation, particularly by FRANK[4] in 1949, gave impetus to similar observations on quartz. Growth patterns were observed on natural quartz by WEIL[5] and WILLIS[6]. A large variety of topographical patterns were found in synthetic quartz, some of which are considered in this discussion.

The present study was carried out in connection with the development† of methods for growing synthetic quartz of low internal mechanical losses. The observations and conclusions presented here are somewhat fragmentary but are considered to be of interest to the art and science of crystal growth. In agreement with investigations on other crystals the hillocks are believed to indicate growth by means of spiral dislocations.

After removal from an autoclave, the grown crystals exhibit well-formed prism and rhombohedral faces. Trapezohedron faces are suggested by narrow planes, which appear as hardly more than wide lines on prism faces.

Even with the unaided eye, considerable topography can be noticed: prisms (10·0) have generally a shingle-like marking, rhombohedral faces show flattened cones, the slower-growing major-*R* faces (10·1) are typically flatter than the others.[7] If a basal pinacoid is present (00·1), it shows a different type of mound and a relatively large number of them.[8] On minor-*r* faces spiral patterns were observed.[9,10]

The crystals were examined with a light-microscope or a metallographic microscope; the

* Paper presented by FRANK AUGUSTINE at *Symposium on Defect Structure of Quartz and Glassy Silica*, Mellon Institute, Pittsburgh, Pa., May 23, 1957. Derived from work performed under a Signal Corps Contract.

† U.S.A. Signal Research and Development Laboratory, Ft. Monmouth, New Jersey, Contract Da-36-039-SC-72415.

VOL.
13
1960

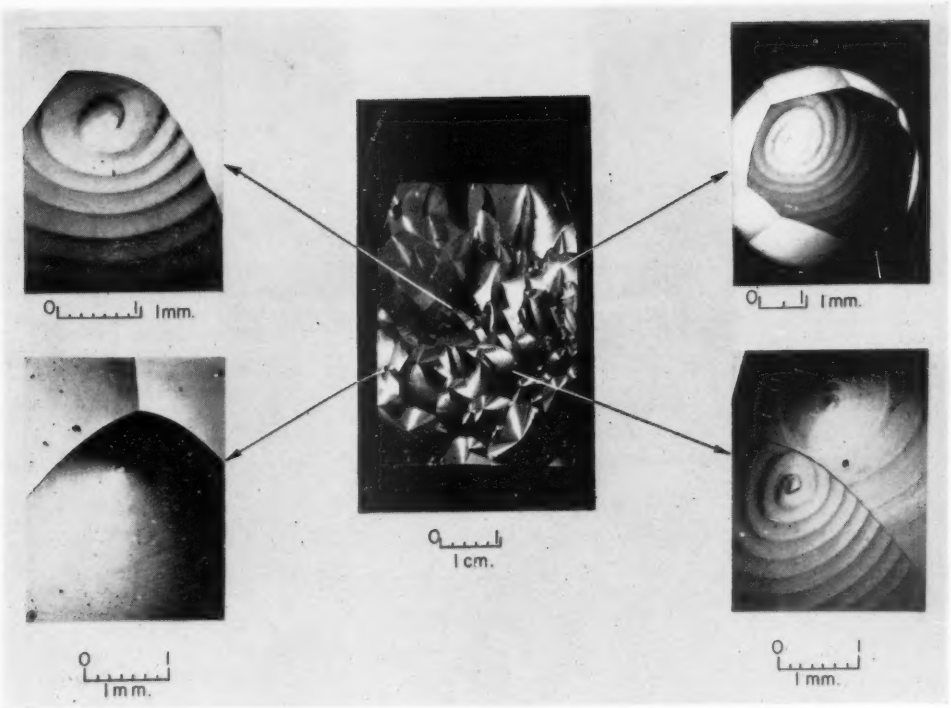


FIG. 1. Minor-*r* face (01·1) composed of many growth hillocks, of which under low magnification seven show definite spiral pattern, in this area of 12 cm².

[*facing p.* 344

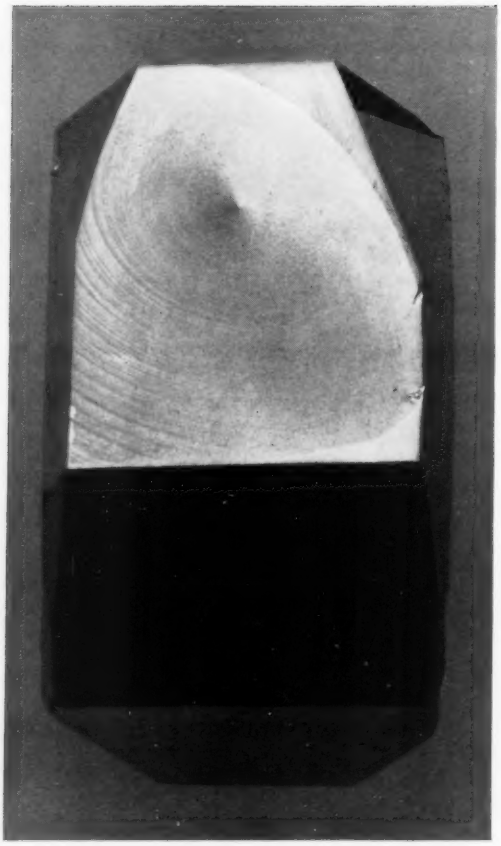


FIG. 2. Minor-*r* face (01·1) displaying a single large hillock on an area of 10 cm². The magnification is 1·8.

VOL.
13
1960

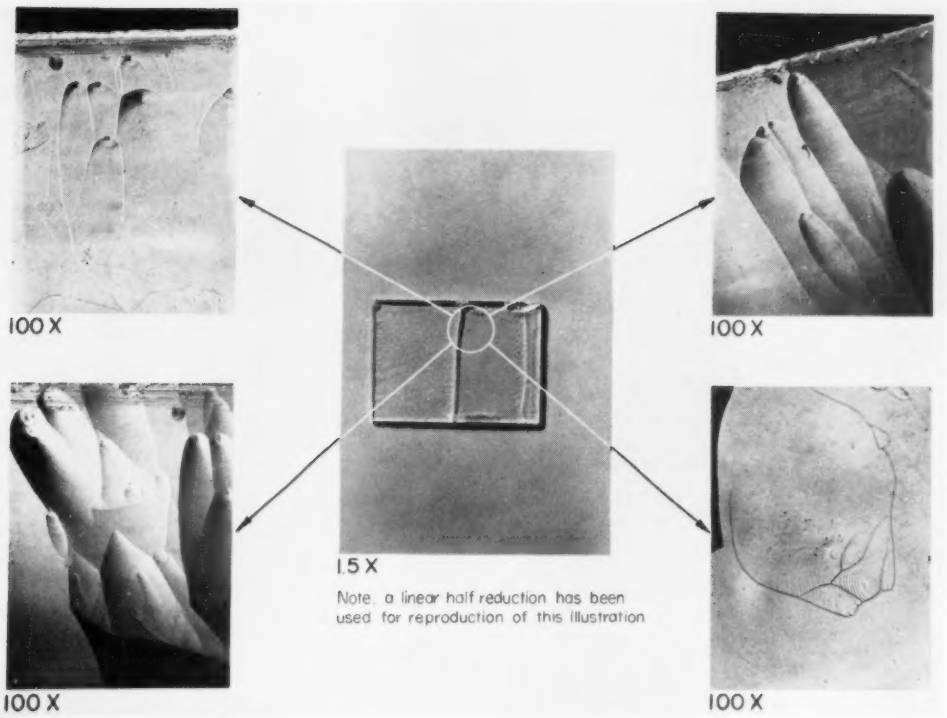


FIG. 3. Minor-*r* face after only 3 days growth, showing a number of hillocks apparently of spiral character.

FIG. 4. Growth spiral of three-fold symmetry on basal pinacoid or (0001) surface. The magnification is 30.

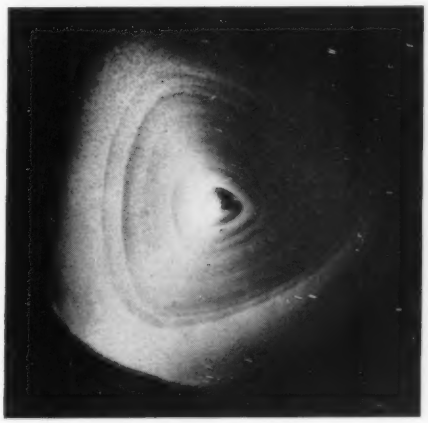


FIG. 5. Terraced cone of three-fold symmetry on (0001) surface. The magnification is 33.

[*facing fig.* 2 (*p.* 344

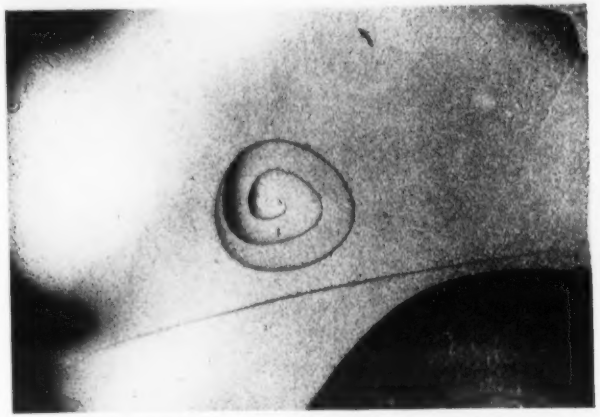


FIG. 6. Growth spiral of two and a half turns on (0001) surface. The magnification is 35.

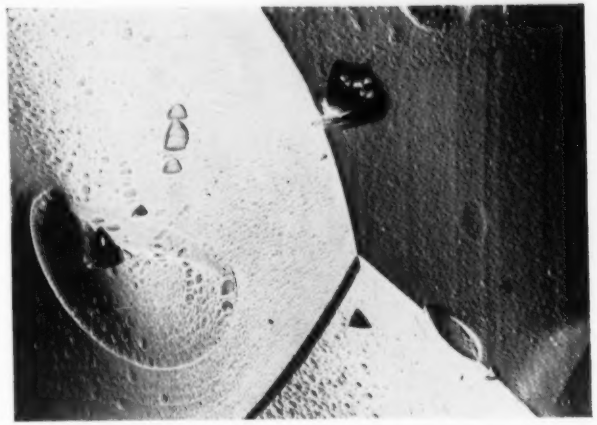


FIG. 7. Etched (01·1) face. The apex of a growth hillock is marked with a deep pit. The magnification is 128.

FIG. 8. Etched (10·1) major-*R* face. In addition to individual pits occurring apparently at random, two coarse lines were developed by the etching process. The magnification is 77.

FIG. 9. Hillock patterns on prism face (10·0). The magnification is 75.

VOL.
13
1960

FIG. 10. Area of Fig. 9 after etching. Each hillock has now been transformed into a pit. The magnification is 69.

FIG. 11. Deep etch pits in (01·1) face as if seen from interior of the crystal. The magnification is 14.

largest magnification used was 230× in the final, enlarged print. It seems probable that interference methods would disclose further interesting details.

Fig. 1 shows a minor-*r* face (01·1) in which 66 well-developed growth hillocks occupy an area of 12 cm². Through the use of evaporated silver and oblique light they are given the appearance of considerable height, although actually the total range in elevation is not more than 0·01 mm. The peaks point in the general direction of the *c*-axis. Seven of the hillocks show a spiral ramp up to a summit; four others suggest a system of spiral terraces.

A single mound, by the way of contrast, covers most of the 10 cm² area of a minor-*r* face shown in Fig. 2. Suggestions of a spiral pattern indicate that a more refined technique of observation would expose a spiral around the obvious summit.

A very short growth time seems to lead to elongated mounds. The minor-*r* seed in Fig. 3 was in contact with supersaturated solution for less than 3 days. It may be assumed that the dislocation line lies more or less centrally in the indicated surface of revolution and thus at a relatively large angle from the normal to the surface.

The synthetic *Y*-bar crystal[11] is characterized by rather imperfect *Z*-surfaces (00·1) and *X*-surfaces (21·0). The *Z*-surface has a pebbled or orange-peel appearance[12] and occasionally the hillocks show spiral patterns. Fig. 4 shows a characteristic spiral of three-fold symmetry, a triangle with rounded points and sides. The bisectors of the suggested triangle, or the three lines of most gentle slope descending from the summit of the cone, are nearly parallel to the *a*-axes of the crystal and are so identified by CINOBER[13]. WILLIAM COOK of this laboratory observed, however, that these lines lie a few degrees to the left of the *a*-axis in left-handed quartz and a few degrees to the right in right-handed quartz.

The speckled background of these three figures seems, on closer examination, to consist of very tiny hillocks subordinated to the high ones. The pattern is surprisingly similar to the spiral terrace pit in germanium metal shown by ELLIS[14].

The pattern in Fig. 5 may, in microscopic form, show evidence of being a spiral ramp, but the appearance here is that of a series of round terraces. The triangularity has the same orientation with respect to the *a*-axes as in Fig. 4.

The solitary snail of a spiral on the *Z*-surface in Fig. 6 might be characterized as an embryo. It seems probable that a new and more active screw dislocation is just getting under way on the gentle slope of a much older cone.

It has been assumed that the growth hillocks are caused by screw dislocations and therefore the summit of each would develop a pit if the quartz were suitably etched. An instance of this is presented in Fig. 7, the etchant being saturated aqueous ammonium bifluoride. The faint outline of several growth steps is seen although the general surface is particularly smooth.

Major-*R* faces occasionally show mounds rather similar to those on the minor-*r* faces. Examples are not shown here. In Fig. 8 the etched surface of an apparently smooth major-*R* face is gouged by numerous triangular pits and two large lines having the appearance of closely coiled springs. These lines lie about 45° with respect to an *a*-axis. On other major-*R* faces rather similar etched lines have been found which are much smaller but again run substantially parallel and about 45° to an *a*-axis. Ammonium bifluoride was again the etchant.

Characteristic hillocks on a prism face are shown in Fig. 9. There appear to be four types of surface features in addition to the fine parallel lines that run toward the upper left corner of the figure. These are:

(1) Rows of distinct hillocks, condensed or spaced, and exemplified particularly by the line extending from the left edge.

(2) Isolated hillocks, each sloping gently toward the upper right.

(3) Shaded areas above each of the isolated hillocks and above some of the hillocks in the lines.

(4) Broken lines of very tiny summits (or even possibly pits) seen along the lower part and upper-left of the picture.

In Fig. 10, hot aqueous alkali etching of the sample of Fig. 9 has produced a pit at each sharp hillock. Nearly each pit in this figure can be seen to have come from a hillock. Most pits are characterized by a dark hole, approximately polygonal, and a shallow-slope pointing toward the lower-right. A second type of pit, believed to be unrelated to dislocations, is very broad and shallow, and squared off at the upper-left. The extensive slopes attached to the hillocks, item 2 of Fig. 9, have left no trace, and the shaded structures, item

VOL. 13 1960

3 of Fig. 9, are faintly observable. The counted density of pits is $4{\cdot}5 \times 10^4$ cm^{-2}.

Fig. 11 presents a view of pits produced by the action of hot aqueous alkaline solution on minor-*r* quartz. The crystal has been sectioned normal to an *a*-axis; this surface has been polished and serves as the "window" through which the rear of the minor-*r* face has been photographed. The undulating surface is part of the rhombohedral face that did not yield pits. The etch pits run deep into the quartz. Perhaps a decorating technique would show that they extend through the crystal. Such a behavior is suggested by the work of DASH[15] on germanium. The pits in general make a small angle with the normal to the minor-*r* face.

Seed plates often become pierced with pits during the heating-up of the pressure vessel. Then, covered by the new growth, they form liquid-vapor inclusions in the harvested crystal. The favored dissolution of the seed at these points indicates that they were probably centers of dislocations. Further, when the grown crystal is deeply etched the pits may point to, or even enter, the liquid-vapor inclusions.

ZIMONYI[16] published in 1957 a number of photographs showing important similarities with certain of the figures presented here. He examined natural quartz and a small number of synthetic crystals. An observation of particular interest is that a large growth cone on a rhombohedral face displayed spiralling lines *after* etching.

In another 1957 publication DESPUJOLS[17] presented photographs of three synthetic crystals showing mound structures, probably spirals, on minor-*r* faces, and discussed etch data and X-ray observations relating to the mosaic structure of quartz.

Acknowledgements are due to Mrs. MARJORIE ASHE, Metalographer, Mechanical Research Division, for the photomicrographs.

REFERENCES

1. HALE D. R. and CHARBONNET W. H., *Conv. Rec. Inst. Radio Engrs. N.Y.* part 3, p. 17 (1954).
2. WALKER A. C. and BUEHLER E., *Industr. Engng. Chem.* **42,** 1369 (1950).
3. BROWN C. S., KELL R. C., THOMAS L. A., WOOSTER N. and WOOSTER W. A., *Miner. Mag.* **29,** 858 (1952).
4. FRANK F. C., *Disc. Faraday Soc.*, no. 5, p. 48 (1949).
5. WEIL A. R., *C.R. Acad. Sci., Paris* **235,** 256 (1952).
6. WILLIS B. T. M., *Nature, Lond.* **170,** 115 (1952); Thesis, University of London, March 1952.
7. HALE D. R., *Rep. 4th Frequency Control Symp. U.S. Army Signal Corps, May* 1952.
8. HALE D. R. and HURLBURT C. S., JR., *Amer. Min.* **34,** 596 (1949).
9. JOST J. M., observations and photography in *6th Quart. Prog. Rep., U.S. Army Signal Corps,* Contract DA-36-039-SC-15381 (1953).
10. HALE D. R. and JOST J. M., *Screw Dislocations Growth Patterns on Synthetic Quartz Crystals,* Amer. Cryst. Assn. Meeting, Cambridge, Mass. April 1954.
11. JAFFE H. and TUROBINSKI T. J., U.S. Pat. 2923604.
12. HALE D. R., *Disc. Faraday Soc.*, no. 5, p. 189 (1949).
13. CINOBER L. I., *Kristallographyia* **1,** 609 (1956).
14. ELLIS S. G., *J. Appl. Phys.* **26,** 1140 (1955).
15. DASH W. S., *J. Appl. Phys.* **27,** 1193 (1956).
16. ZIMONYI G., *Acta Phys. Hung.* **8,** 119 (1957).
17. DESPUJOLS J., Thesis, 1957, Faculty of Sciences, University of Paris.

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 13. pp. 347–351. Printed in Great Britain.

STRUCTURAL EFFECTS IN THE CHEMICAL REACTIVITY OF SILICA AND SILICATES*

F. M. ERNSBERGER

Glass Research Center, Pittsburgh Plate Glass Company, Harmarville, Pennsylvania

(*Received* 27 *May* 1959)

VOL. 13 1960

Abstract—The attack of quartz by hydrofluoric acid proceeds at a rate approximately a hundredfold greater on the optic axial plane than on the prism faces. The rate of attack on coesite is nil. Mica and certain other silicates are resistant to hydrofluoric acid attack on certain crystal faces. A reaction mechanism is proposed to account for these facts. This mechanism predicts that in a perfect crystal, only those crystal faces which expose simultaneously two of the four oxygen atoms of a given silica tetrahedron are attacked by hydrofluoric acid. Coesite is a special case, in that the close-packing which is assumed to exist provides double the normal number of tetrahedral sites. Surface silicon atoms may occupy some of these vacant sites, with the result that no unsaturated oxygen atoms are exposed at the surface.

INTRODUCTION

It is a well-known generalization of chemistry that silica and the silicates are chemically attacked by aqueous hydrofluoric acid. This statement oversimplifies a rather complicated situation, as a few examples will indicate.

Alpha quartz is attacked at a rate at least a hundredfold greater on a plane perpendicular to the optic axis than on the prism faces. The effect is conveniently demonstrated by etching a sphere prepared from a single crystal.[1] Mica also shows a remarkable directional effect. The cleavage surface is so resistant to aqueous hydrofluoric acid that mica has been used successfully as a window material for microscopic study of the etching of glass. It has been reported that "dense silica" (coesite) is not attacked at all by hydrofluoric acid.[2]

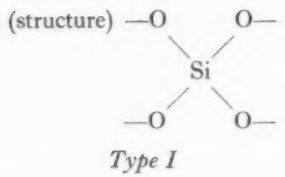
These anomalies are of a magnitude entirely too great to be explained by minor influences such as surface energy or packing density variations. One must look instead for some influence which is capable of completely blocking the chemical reaction in a perfect crystal. In this paper the viewpoint is adopted that surface structure holds the key: that certain two-dimensional arrangements of SiOSi bonds are inherently resistant to chemical attack.

* Paper presented at *Symposium on Defect Structure of Quartz and Glassy Silica*, Mellon Institute, May 22–24, 1957.

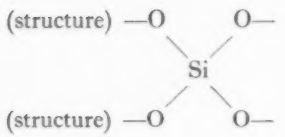
SURFACE STRUCTURE OF SILICATES

Current theories of the structure of silica and silicates recognize the tetrahedral arrangement of four oxygen atoms around one silicon atom as a common feature of nearly all silicates, both crystalline and vitreous. An elementary fact which appears to go unrecognized is that there are only four ways in which this structural element can be exposed at the surface of a solid. These ways are distinguished by the number of oxygen atoms remaining attached to the subsurface structure (Fig. 1).

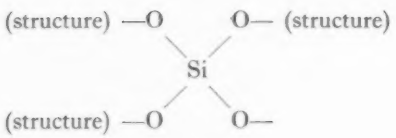
According to this concept, it should be possible to classify any silicate surface as one or a combination of these elementary exposure types. In vitreous phases, for example, the four types will occur at random. They need not occur in equal numbers, of course. Type I, for example, is undoubtedly too unstable to have any prolonged existence in the presence of water. In crystals, on the other hand, the exposure type will depend on the crystal plane or planes which constitute the surface.



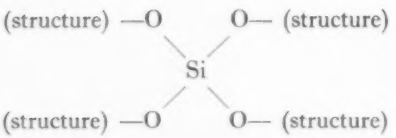
Type I



Type II



Type III

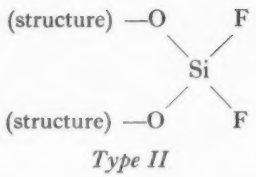


Type IV

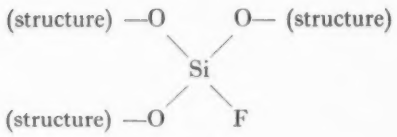
FIG. 1. The four possible surface exposures of the silica tetrahedron.

SURFACE REACTIONS

In aqueous systems the exposed oxygen valencies of Fig. 1 will ordinarily be electrically neutralized by protons, forming a surface layer of hydroxyl groups. In the presence of fluoride ions, however, there is the alternative possibility that the silicon ion will coordinate fluoride ions. This substitution is encountered quite frequently in crystal structures, owing to the similarity in size and charge between OH^- and F^-. An extreme example of surface substitution is observed when glass is immersed in concentrated sulfuric acid to which a trace of a fluoride has been added. The surface becomes so nonpolar that it ceases to be wetted by the acid.[5] It may be assumed that a significant equilibrium proportion of "hybrid" tetrahedra (Fig. 2)



Type II

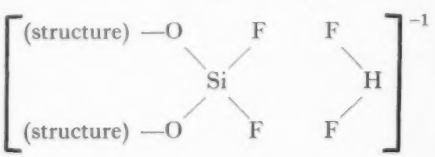


Type III

FIG. 2. "Hybrid" tetrahedra resulting from the replacement of OH^- by F^-.

exists under less extreme conditions of acidity. It is probable that some substitution occurs even in the neighborhood of pH = 7, where there is no observable chemical attack.

The coordination of fluoride ions at the surface in the manner illustrated in Fig. 2 satisfies the electrical neutrality requirement, but leaves the silicon atom with a serious deficiency in its coordination, which is known to be six with respect to fluorine, rather than four. Therefore the Type-II hybrid tetrahedron will tend to coordinate a further complement of small negative ions. This may lead to the adsorption of further fluoride ions, but there should be a particular preference for bifluoride ions. This species has the unique property of presenting simultaneously a pair of fluoride ions in contact. Thus the bifluoride ion is able to fill simultaneously two coordination sites in the manner of a chelating agent, with a corresponding increase in stability. The absorption of a bifluoride ion by a Type-II surface would lead to the formation of a complex having the configuration



Granting the possible existence of such a surface complex, it is easy to account for continuous chemical reaction at a Type-II surface. The four

fluorine atoms in the immediate vicinity of the surface silicon provide a second tetrahedral coordination site as an alternative to the one currently occupied by the silicon atom. A moderate degree of thermal activation should suffice to shift the silicon atom from one site to the other, with formation of SiF_4. The ready availability of protons may be expected to ease the transition by occupying the oxygen valencies vacated by the silicon atom. The newly-formed SiF_4 molecule will immediately coordinate two more fluoride ions to form the fluorosilicate ion. It may in fact do so as part of the transition process.

The Type-III surface, on the other hand, is clearly unreactive. Steric considerations make it impossible to construct an alternative coordination site in close proximity to the silicon atom. Moreover, the atom in question is bound to the subsurface structure by three linkages instead of two. The Type-IV surface, for the same reasons, should be completely unreactive.

This attack mechanism is in qualitative agreement with known facts about the reaction of silicates with aqueous fluorides. In particular, it accounts for the fact that only acidified solutions lead to measurable attack, and that attack rate goes through a maximum with bifluoride ion concentration.[3,4]

CONCLUSIONS

Directional effects in the dissolution of quartz are satisfactorily explained by the fact that the rapidly-attacked basal plane presents a Type-II exposure, while the prism faces are Type-III. These relationships are illustrated in Fig. 3 for the case of beta quartz. In the case of alpha quartz the distribution of bonds is identical, though a puckering of the lattice exists which reduces the

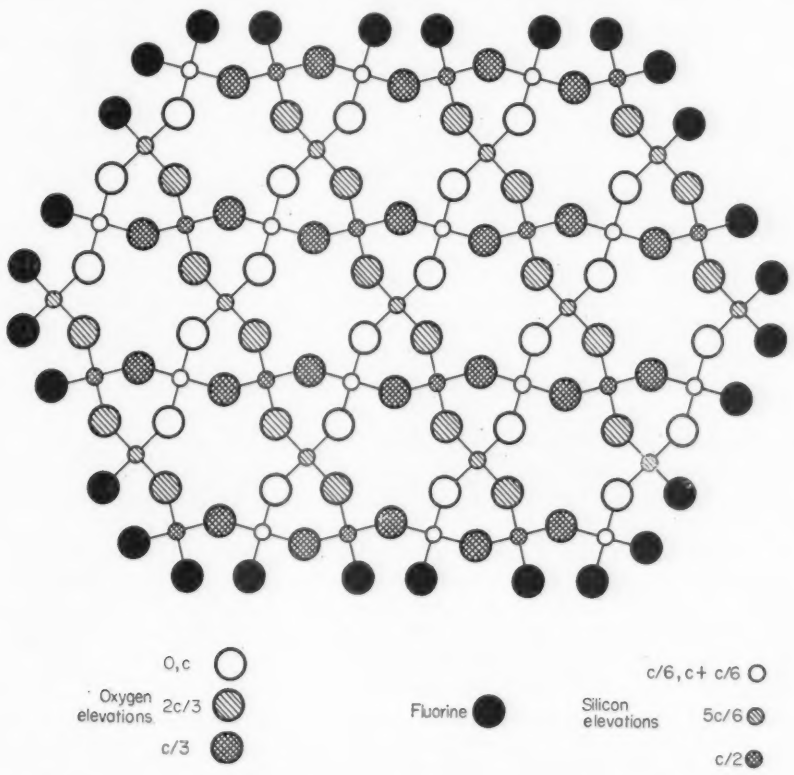


FIG. 3. The structure of beta quartz, projected on a plane perpendicular to the optic axis. Type-III prism-face terminations are shown occupied by fluorine atoms.

symmetry from hexagonal to trigonal. It will be observed that prism-face exposures in successive layers of atoms alternate between Types II and III. In the presence of chemical reagents, the surface will necessarily consist of the chemically-resistant Type-III, as in Fig. 3.

In the case of mica, three of the four vertices of each silica tetrahedron are occupied in forming hexagonal silicon–oxygen sheets. This would lead to a Type-III structure, were it not that these sheets occur in firmly-bound pairs with the fourth vertex of each tetrahedron sandwiched between the sheets, where it is not available for substitution. The cleavage planes in the resulting structure are therefore Type-IV, and consequently unreactive. This unreactivity is not true of the edges of the sheets, of course, so it would be expected that finely-ground mica would be soluble in hydrofluoric acid. The same considerations apply to other layer-structured minerals related to mica (talc, pyrophyllite, etc.) which occur only as fine powders.

The interesting case of coesite cannot be immediately fitted into the picture because the detailed structure is still unknown. The available information suggests that the unreactivity is to be ascribed to the existence of a Type-IV surface.

The high density of coesite (3·01) requires an extremely compact arrangement of oxygen atoms. In order to account for the observed density it is necessary to assume closest packing (face-centered cubic or hexagonal) of oxygen ions of radius 1·3 Å. X-ray and petrographic studies indicate that the symmetry is not this high,[2,6] but for purposes of the present discussion it will be assumed that coesite approaches the face-centered cubic structure. This is in agreement with the observation[2] that coesite reverts at high temperature into cristobalite.

The structural relationship between beta cristobalite and the hypothetical close-packed coesite is shown in Fig. 4. The effect of high pressure and temperature is seen to be a collapse of the relatively open structure of cristobalite, but without any rearrangement of bonding partners. Fig. 4 also illustrates a fact significant to the present discussion; namely, that the formation of a close-packed structure increases the number of tetrahedral sites suitable for occupation by a silicon ion. In cristobalite every such site is occupied, but in the close-packed structure there are four times as many tetrahedral sites as silicon ions. Under these circumstances, a silicon ion exposed at the surface (represented in Fig. 4 by half-spheres) may complete its coordination by an alternative to the usual method of coordinating hydroxyl groups and water molecules: it can drop beneath the surface into one of the vacant tetrahedral sites (Fig. 4c). This migration of surface silicon ions will be accompanied by a polarization of the equivalent number of oxygen bonds toward the subsurface silicon ions, leaving the surface saturated and unreactive.

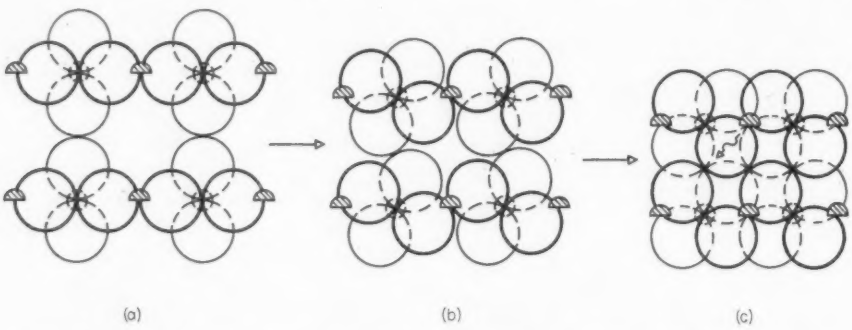


FIG. 4. The relationship between beta cristobalite and an idealized coesite. (a) beta cristobalite; (b) intermediate structure; (c) coesite. Surface silicon atoms are represented as half-spheres to denote the fact that half the silicon ion sites must be vacant in a cleavage surface.

It is interesting to speculate whether other close-packed oxygen lattice structures might be expected to show anomalous solution kinetics. The mineral bromellite (BeO) has the close-packed structure,

with Be ions occupying half the existing tetrahedral interstices. Thus there are subsurface sites available, but the smaller charge of the beryllium ion would probably make surface polarization effects negligible. Corundum (Al_2O_3) is also close-packed but the aluminum ions occupy octahedral rather than tetrahedral sites. Hence, there are not enough empty sites to accommodate all the surface aluminum ions in the atomic layer immediately below the surface. In general, any thermodynamically stable structure such as those mentioned would not be expected to have as strong a polarizing influence on its surface cations as does the metastable coesite. The latent energy associated with the creation of the large number of unoccupied tetrahedral sites in coesite must render these sites extraordinarily deep potential wells with respect to occupation by a silicon ion.

VOL.
13
1960

REFERENCES

1. MUGGE O., *Rosenbusch Festschrift* (1906); quoted in *Amer. Min.* **30,** 315 (1945). The effect has been confirmed in the author's laboratory.
2. COES L. Jr., *Science* **118,** 131 (1953).
3. PALMER W. G., *J. Chem. Soc.* p. 1656 (1930).
4. ROTHMAN A. B., Pittsburgh Plate Glass Research Laboratories, unpublished data.
5. SIMONS J. H., *Fluorine Chemistry* Vol. I, p. 570. Academic Press, New York (1950).
6. RAMSDELL L. S., *Amer. Min.* **40,** 975 (1955).

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 13. p. 352. Printed in Great Britain.

THE ANELASTICITY OF NATURAL AND SYNTHETIC QUARTZ AT LOW TEMPERATURES—ABSTRACT*

J. C. KING

Certain imperfections in alpha-quartz are found to introduce anelastic absorption at low temperatures in crystal resonators. For a shear wave of 5 mc, absorption peaks are observed at approximately 20° and 50°K in unirradiated samples. X-irradiation effects a lowering of the amplitude of the 50°K absorption, while a new absorption is introduced at 100°K.

The results of acoustic absorption measurements on a number of selected samples of synthetic quartz resonators suggest that, whereas the 50°K absorption is attributable to excess oxygen, the 100°K defect involves aluminum which is present substitutionally in the crystal structure.

X-irradiation-induced lowering of the resonant frequency of crystal resonators is clearly the result of a lowering of the elastic modulus associated with the 100°K anelastic absorption.

* Paper presented at *Symposium on Defect Structure of Quartz and Glassy Silica*, Mellon Institute, May 22–24, 1957. Published in *Bell Syst. Tech. J.* **38**, 573–603 (1959).

LETTER TO THE EDITOR

Sur la Résistivité et le déplacement de Knight d'alliages liquides de sodium

(*Received* 11 *September* 1959)

VOL. 13 1960

1. *Introduction*

SUBSTITUER un atome étranger B à un atome normal dans un métal A perturbe la structure électronique de ce métal. Des considérations théoriques permettent, dans le cas de l'état solide, de ramener ce problème à celui de la diffusion des électrons de conductibilité de A, traités comme des électrons libres, par un potentiel en puits carré représentant l'atome B dissous.[1, 2] Ce modèle simple a donné de bons résultats pour le calcul des résistivités résiduelles[3] et des déplacements de Knight[2, 4] des solutions solides diluées. L'étude suivante montre qu'il en est de même, à l'état liquide, pour les alliages du sodium avec les autres métaux alcalins, tant en ce qui concerne la résistivité que le déplacement de Knight du sodium. Par contre, cette approximation n'est plus bonne quand l'électronégativité du métal dissous diffère nettement de celle du sodium: ainsi pour l'argent, l'or et le mercure.

2. *Modèle utilisé*

Nous supposons que la dissolution de B dans A (ici, le sodium), se fait sans contraction importante. Nous prenons alors pour rayon du puits de potentiel, représentant l'atome B dissous, le rayon atomique du métal B pur. La profondeur du puits est déterminée par approximations successives de façon que le gaz électronique perturbé fournisse la charge d'écran convenable à l'atome B dissous. Selon BLATT[3] cette charge vaut :

$$Q = Z + \frac{V_A - V_B}{V_A} p$$

fois celle de l'électron, Z étant la différence de valence entre A et B (nulle pour les alcalins), V_A et V_B les volumes atomiques respectifs de A et de B et p le nombre d'électrons de conductibilité par atome A. Si on analyse en composantes sphériques autour de B la fonction d'onde d'un électron du niveau de Fermi de A, sa diffusion par B introduit un déphasage η_l dans la lème onde partielle. Ces déphasages sont liés à la charge d'écran par la relation[5] :

$$Q = \frac{2}{\pi} \sum_l (2l+1)\eta_l$$

Ceci permet de déterminer la profondeur correcte du puits de potentiel; pratiquement, il suffit souvent de se limiter aux trois premiers déphasages.

3. *Calcul de la résistivité*

La résistivité résiduelle $\Delta\rho$ due à la concentration c d'atomes B dissous dans A s'exprime en unités atomiques ($e = m = \hbar = 1$), par la relation[6] :

$$\Delta\rho = \frac{4\pi c}{kp} \sum_l \sin^2(\eta_{l-1} - \eta_l),$$

k désignant le rayon de la sphere de Fermi du métal A. Nous admettons que cette formule donne encore correctement, à la température de fusion du sodium, l'accroissement de résistivité dû aux atomes dissous, par rapport à celle du sodium pur.

Le tableau 1 permet de comparer les valeurs calculées de cette façon aux résultats expérimentaux récemment obtenus par FREEDMAN et ROBERTSON[7] pour une concentration atomique c de 1 pour cent :

Tableau 1. *Comparaison des accroissements de résistivité calculés et mesurés* ($c = 1$ *pour cent*)

Alliage	$\Delta\rho_{calc}$($\mu\Omega$-cm)	$\Delta\rho_{mes}$($\mu\Omega$-cm)
Na–Li	0,43	0,3
Na–K	1,15	1,1
Na–Rb	2,15	2,5
Na–Cs	3,7	3,75
Na–Ag	1,2	2,7
Na–Au	1,2	5

On remarque un bon accord pour les éléments alcalins, mais pas pour l'argent ni l'or. Cette différence de comportement se confirme dans l'étude du déplacement de Knight du sodium contenant ces atomes en solution.

4. *Calcul du déplacement de Knight du sodium*

On sait que la fréquence ν de résonance magnétique nucléaire d'un métal pur est plus élevée que si ce métal se trouve combiné dans un sel. Le déplacement relatif : $K = \Delta\nu/\nu$ est proportionnel à la densité aux noyaux des électrons du niveau de Fermi.

Dans l'alliage $A\,B$, la perturbation de la densité électronique de A par les atomes dissous B modifie le déplacement de Knight K du métal A. Le changement relatif $\Delta K/K$ est proportionnel à la concentration c en atomes B, au moins tant que celle-ci n'est pas trop forte :[4]

$$\frac{\Delta K}{K} = c \sum_l (\alpha_l \sin^2\eta_l + \beta_l \sin 2\eta_l)$$

$$\text{avec} \begin{cases} \alpha_l = (2l+1) \sum_R [n_l^2(kR) - j_l^2(kR)] \\ \beta_l = -(2l+1) \sum_R j_l(kR) n_l(kR) \end{cases}$$

Dans ces formules, j_l et n_l sont respectivement les fonctions sphériques de Bessel et de Neumann; $\boldsymbol{R}$ désigne le rayon vecteur d'un noyau A à partir du noyau B pris pour origine.

Dans un alliage liquide, près du point de fusion, nous admettons qu'il subsiste un ordre local autour de chaque atome dissous. Nous supposons que chacun d'eux est entouré de huit premiers voisins, comme dans la structure cubique centrée des métaux alcalins; ou bien, quand l'atome dissous est beaucoup plus gros que le sodium, nous envisagerons la possibilité d'un environnement compact de douze premiers voisins, comme dans la structure cubique à faces centrées. A l'extérieur du volume occupé par ces atomes, nous remplaçons la somme sur les sites R par une intégrale de volume, de façon à tenir compte de la répartition aléatoire des noyaux dans le liquide.

Le tableau 2 permet de comparer les valeurs ainsi calculées aux résultats expérimentaux de Bloembergen et Rimaï[8], d'une part d'Oriani et Webb[9], d'autre part :

L'accord entre les valeurs mesurées et calculées est assez bon quand les deux constituants de l'alliage sont des métaux alcalins; il est beaucoup moins bon pour l'or et le mercure dans le sodium. Ceci est à rapprocher du fait que la dissolution du mercure dans le sodium se fait avec une contraction considérable.[10] Dans ces conditions, il n'est plus justifié de prendre pour rayon du potentiel diffuseur le rayon atomique du mercure pur. Cette contraction indique que la perturbation affecte fortement les atomes de sodium voisins du mercure dissous; notre modèle n'est plus valable dans ce cas.

Tableau 2. Variations par unité de concentration du déplacement de Knight du solvant dans les alliages liquides.

Alliage	$1/C\ \Delta K/K$		
	Calculé		Mesuré
	C.C.*	C.F.C.†	
Na–Li	−0,57		−0,5 [9]
Na–K	0,1	0,25	0,35[8]
Na–Rb	0,15	0,4	0,5 [8]
Na–Cs		0,67	0,7 [9]
Rb–Na	−0,23		−0,25[8]
Cs–Na	−0,3		−0,43[9]
Na–Au	−0,2		−0,5 [9]
Na–Hg	−0,1		−0,5 [9]

* C.C. : Environnement des premiers voisins formant une structure cubique centrée.

† C.F.C. : Environnement des premiers voisins formant une structure cubique à faces centrées.

5. *Conclusions*

Nous avons appliqué à des alliages liquides d'éléments de même valence, un modèle conçu initialement pour les solutions solides diluées d'éléments de valences différentes. Ce modèle nous a donné simultanément des valeurs de la résistivité et du déplacement de Knight en assez bon accord avec l'expérience pour les alliages formés de métaux alcalins. Il cesse de s'appliquer quand l'électronégativité du soluté diffère notablement de celle du solvant; ceci suggère de considérer ensemble, dans ce cas, l'atome dissous et ses plus proches voisins, dont les nuages électroniques doivent être fortement perturbés.

Remerciements—Nous remercions les Docteurs R. A. Oriani, L. Rimaï et W. D. Robertson de nous avoir

aimablement communiqué leurs résultats expérimentaux avant publication.

Physique des Solides E. DANIEL
60 *Boulevard St. Michel*
Paris (6e).

Bibliographie

1. DANIEL E., *J. Phys. Chem. Solids* **10,** 174 (1959).
2. DANIEL E., Thèse, Paris (1959).
3. BLATT F. J., *Phys. Rev.* **108**, 285 (1957).
4. BLANDIN A., DANIEL E. et FRIEDEL J., *Phil. Mag.* **4**, 180 (1959); BLANDIN A. et DANIEL E., *J. Phys. Chem. Solids* **10,** 126 (1959).
5. FRIEDEL J., *Advances in Phys.* **3**, 446 (1954).
6. DE FAGET DE CASTELJAU P. et FRIEDEL J., *J. Phys. Radium* **17**, 27 (1956).
7. FREEDMAN J. F. et ROBERTSON W. D., communication privée (1959).
8. BLOEMBERGEN N. et RIMAÏ L., communication privée (1958).
9. ORIANI R. A. et WEBB M. B., *Acta Met.* **7**, 63 (1959); communication privée (1959).
10. VANSTONE E., *Trans. Faraday Soc.* **7**, 42 (1911).

VOL.
13
1960

NOTICE

The Sixth Annual Conference on Magnetism and Magnetic Materials will be held in New York City, November 14–17, at the New Yorker Hotel. This conference is sponsored jointly by the American Institute of Electrical Engineers and the American Institute of Physics, in co-operation with the Office of Naval Research, the Institute of Radio Engineers and the Metallurgical Society of the A.I.M.E.

Authors should submit titles and abstracts of proposed papers by 26 August to A. M. CLOGSTON or R. C. FLETCHER, Program Chairman, Bell Telephone Laboratories, Murray Hill, New Jersey. Further conference details can be obtained from the Local Chairman, L. R. BICKFORD, JR., IBM Research Center, Yorktown Heights, New York.

ERRATUM

R. F. BREBRICK, Minimum molar Gibbs free energy and small deviations from stoichiometry in crystalline binary compounds, *J. Phys. Chem. Solids* **11**, 43 (1959).

The energies to create a cation and anion vacancy have been incorrectly given as ζ_{hC} and ζ_{hA} respectively. Inspection of the Boltzman factor in the grand partition function[5] shows that the energy required to remove an X-atom from the crystal to the standard state leaving an anion vacancy in the crystal and an electron at the Fermi level, the total concentration of cation vacancies being held fixed, is given by $\zeta_{hA}-U+E_f$. The energy to create a cation vacancy under analogous conditions is $\zeta_{hC}-U-E_f$. Removing the unnecessary restriction that the crystal be simultaneously stoichiometric and intrinsic, the term $\zeta_{hC}-\zeta_{hA}$ in the right-hand member of equation (6) becomes $\zeta_{hC}-\zeta_{hA}-2E_f^i$ where E_f^i is the Fermi level in the intrinsic crystal. This quantity then is the difference in the energy to create a cation vacancy in the intrinsic crystal and that to create an anion vacancy.

BOOKS RECEIVED

E. GRUNWALD and R. H. JOHNSEN: *Atoms, Molecules and Chemical Change*. Prentice-Hall Inc., New York, 1960. 252 pp., $6.00.

Syntheses of Heterocyclic Compounds, Volumes 1 and 2. Edited by A. L. MNDZHOIAN. Consultants Bureau, Inc., New York, 1959. 84 pp., 48s.

L. DE BROGLIE: *Non-Linear Wave Mechanics—A Causal Interpretation.* Elsevier Publishing Company, Amsterdam, 1960. 304 pp., 57s.

Annual Reports on the Progress of Chemistry, Volume LV, 1958. The Chemical Society, London, 1959. 527 pp., £2 . 0 . 0.

VOL.
13
1960